Aalto University
School of Electrical Engineering
Degree Programme in Communications Engineering

Juho Rinne

Mechanical Distorters in Musical Stringed Instruments

Master's Thesis
Espoo, November 19, 2014

Supervisor: Professor Vesa Välimäki
Advisor: Henna Tahvanainen, M.Sc. (Tech.)

Aalto University
School of Electrical Engineering
Degree Programme in Communications Engineering

ABSTRACT OF MASTER'S THESIS

Author:	Juho Rinne		
Title: Mechanical Distorters in Musical Stringed Instruments			
Date:	November 19, 2014	**Pages:**	63
Major:	Acoustics and Audio Signal Processing	**Code:**	S-89
Supervisor:	Professor Vesa Välimäki		
Advisor:	Henna Tahvanainen, M.Sc. (Tech.)		

Some acoustic stringed instruments have been designed for quiet environments, and they do not produce sound loud enough for today's noisy environment. The human ear perceives high-frequency sounds louder than lower ones. Thus, mechanical distorters have been designed to transfer energy from the low end of the spectrum to the more cutting high end. This thesis studies mechanical distorters, which have been made from curved metal wire and which lightly touch the string near its end point. It describes how three different mechanical distorters affect on the timbre of a vibrating string in a kantele, a traditional Finnish folk instrument. Both distorted and undistorted tones were recorded, and their spectrograms, the temporal evaluation of their partials, and sound pressure levels were compared. According to this study, mechanical distorters transfer energy from the low end of the spectrum to the high end by generating new long-lasting upper harmonics, thus making the sound brighter. Based on comparison of the standard sound pressure level (SPL) and the A-weighted SPL of the tones the distorted strings are perceived louder.

Keywords:	Acoustics, music, kantele, harmonic distortion, nonlinear systems, spectral analysis
Language:	English

Aalto-yliopisto
Sähkötekniikan korkeakoulu
Tietoliikennetekniikan koulutusohjelma

DIPLOMITYÖN TIIVISTELMÄ

Tekijä:	Juho Rinne		
Työn nimi: Mekaaniset säröttimet kielisoittimissa			
Päiväys:	19. marraskuuta 2014	**Sivumäärä:**	63
Pääaine:	Akustiikka ja äänenkäsittely	**Koodi:**	S-89
Valvoja:	Professori Vesa Välimäki		
Ohjaaja:	Diplomi-insinööri Henna Tahvanainen		

Jotkut akustiset kielisoittimet on suunniteltu hiljaiseen ympäristöön, eivätkä ne tuota tarpeeksi ääntä nykyajan meluisampiin ympäristöihin. Koska ihminen aistii korkeataajuuksiset äänet voimakkaampina kuin matalat, mekaaniset säröttimet on tehty siirtämään energiaa alemmilta taajuuksilta korkeammille ja leikkaavammille taajuuksille. Tämä työ tutkii mekaanisia säröttimiä, jotka on tehty taitellusta metallilangasta, joka koskettaa kevyesti kieltä sen päätepisteen lähellä. Työssä kerrotaan, miten kolme eri särötintä vaikuttaa kanteleen värähtelevän kielen äänensävyyn. Sekä säröttämätön ja särötetyt äänet on äänitetty ja niiden spektrogrammeja, yksittäisten harmonisten kehittymistä ja äänipainetasoja on vertailtu. Tämän työn perusteella mekaaniset särötimet siirtävät energiaa alemmilta taajuuksilta ylemmille taajuuksille luoden pitkäkestoisia korkeita harmonisia, tehden näin äänestä kirkkaamman. Särötetty kieli kuulostaa voimakkaammalta, kun verrataan normaalia äänipainetasoa ja A-painotettua äänipainetasoa.

Asiasanat:	Akustiikka, musiikki, kantele, harmoninen särö, epälineaariset järjestelmät
Kieli:	Englanti

Acknowledgements

I would like to thank my supervisor Vesa Välimäki and advisor Henna Tahvanainen for guiding me through this process. Additionaly, thank you to Jyrki Pölkki who designed the studied kantele and the distorters.

I would also like to thank my fiancée, parents, and friends for their support.

Espoo, November 19, 2014

Juho Rinne

Symbols and Abbreviations

Symbols

c	speed of transverse wave
c_0	speed sound in air
I	intensity
I_0	reference intensity
L_I	sound intensity level
L_W	sound power level
p	sound pressure
p_0	reference sound pressure
T	tension
t	time
W	sound power
W_0	reference sound power
x	coordinate on the string
y	displacement of the string
μ	linear mass density
ρ	density of air

Abbreviations

SPL	sound pressure level
T30	time to decay 60 decibels estimated from time to decay 30 decibels
T60	time to decay 60 decibels

Contents

Chapter 1

Introduction

The traditional Western stringed instruments are fairly quiet and their spectral content is focused at the low end of spectrum. They have been designed to quiet environments, such as small rooms or concert halls which are designed to emphasize high frequency content [1]. In a controlled environment everything works well, but in a noisy environment, such as a bar or a busy street, they get masked under the loud background noise.

In order for those instruments to be heard in a noisy environment, their loudness or spectral content has to be modified. The loudness can be increased with electrical amplification but then the acoustic instrument may lose its acoustic qualities. Alternatively, the shape of the body of the instrument can be modified in a way that it radiates the sound waves more efficiently. Penttinen et al. [2] and Tahvanainen [3, 4] studied a modified kantele with increased string tension, an increased radiation area, and an additional air gap between the plates. Instead of only increasing the loudness, the spectral content of the instrument can be modified. Since the human ear perceives high frequencies and broadband tones louder than low frequencies and narrow tones, the loudness of an instrument may be increased by transferring energy from low frequencies to higher frequencies with mechanical distorters.

A mechanical distorter is a piece or surface that is designed to create nonlinearities. The purpose of distorters is to alter the timbre of the instrument and increase its loudness. For example many Eastern stringed instruments, such as the biwa and the sitar, have a curved bridge which produces broadband sounds with emphasized high frequencies. This helps their sound to cut through loud background noise of the busy Indian streets.

Various studies of the interaction between a vibrating string and an obstacle have been presented. Schatzman [5] and Burridge et al. [6] studied the vibrating string constrained by a rigid obstacle. Legge and Fletcher [7] examined theoretically the nonlinear generation of missing modes on vibrating strings. Frontini and Gotussu [8] presented theoretical and experimental results of a string vibrating against an obstacle. Rank and Kubin [9] created a model for slap bass synthesis. Krishnaswamy and Smith [10], and Evangelista and Eckerholm [11], presented methods for simulation of string-obstacle collisions. Taguti [12] created a model for the characteristic buzzing tone of the Indian lutes biwa and shamisen. Lehtonen et al. [13] studied the nonlinear effect of part-pedaling effect in the piano. Stulov and Kartofelev [14] created a model for nonlinear effects of the string vibrations caused by nonlinear supports. Bilbao and Torin [15] studied the collisions between the string, the fretboard and the individual frets of the guitar.

This thesis studies the mechanical distorters that are attached to a kantele, a traditional Finnish instrument which belongs to the zither family. The studied kantele is a fretless instrument that has eleven strings that do not extend beyond its body. The designed mechanical distorters are made of curved metal wires, paperclips, and rawhide. The distorters are placed near the end point of the string so that they almost touch the string. The mechanical distorter limits the amplitude of the string thus altering the timbre of the kantele. The undistorted and distorted tones were recorded and the temporal evolution of their spectrograms, temporal harmonic partials, spectral centroids and signal power levels were compared. This thesis does not pay any special attention to the body of the kantele, since the string vibrations solely cause the audible characteristics of the kantele and the mechanical distorters. According to studies the body of the kantele acts as a resonant high pass filter [3, 16].

This thesis is structured as follows. At first the ideal stringed instruments are being studied. The second chapter is about hearing and describes how human auditory system works. Additionally, it explains what kind of sounds a human perceives louder than the others and justifies the usage of distortion. The third chapter explains what is the distortion and gives some examples of the existing instruments with nonlinearities. The fourth chapter describes the measurement setup of the studied distorters and the fifth chapter shows the results.

Chapter 2

Plucked Stringed Instruments

This chapter describes the features of a plucked stringed instrument. The family of plucked stringed instruments consists of guitars and lutes as well as stringed instruments from Asia, such as shamisen, biwa, and sitar [1]. The timbre of Asian instruments is very different than the timbre of the Western instruments. However, each of them has vibrating strings and one or more soundboard.

A musical string is a very thin object stretched between two points which are attached to the body of the instrument [17]. The vibration of the string transfers to the body of the instruments via a tailpiece called the bridge or via the tuning pegs. The body and the possible air cavity inside work as a resonator and radiate the amplified sound.

2.1 Vibrating String

The vibrating string can be though as a mass-spring system with a very large amount of masses [18]. The string has mass and elasticity. In instruments such as the violin and guitar, the other end of the strings are attached to the tail piece while the other end is attached to the tuning pegs [18]. Plucking of the string creates transverse waves that propagate in both directions across the string [1]. In the case of an ideal string the waves obey the simple wave equation

$$f(x,t) = \frac{\delta^2 y}{\delta t^2} - c^2 \frac{\delta^2 y}{\delta x^2}, \tag{2.1}$$

where the c is the speed of the transverse wave, t is time, x is the coordinate on the string and y is the displacement of the string [19]. $\frac{\delta y}{\delta t}$ and $\frac{\delta y}{\delta x}$ are the temporal and spatial derivatives. For a plucked string the $f(x,t)$ becomes zero after the plucking [20]. If a string with a linear mass density μ (kg/m) is stretched to a tension T, waves will propagate at speed c given by $\sqrt{T/\mu}$. The general solution of the wave equation can be written as

$$y = f_1(ct - x) + f_2(ct + x). \tag{2.2}$$

In the equation the overall displacement y is the sum of individual displacements of waves f_1 and f_2. The function $f_1(ct-x)$ represents waves travelling to the right of the plucking point and the function $f_2(ct+x)$ represents waves travelling to the left with the same velocity. When they are in the same phase the displacement is the sum of both of the waves and when they are in opposite phase they are subtracted from each other. Figure 2.1 shows the superposition of the two travelling waves.

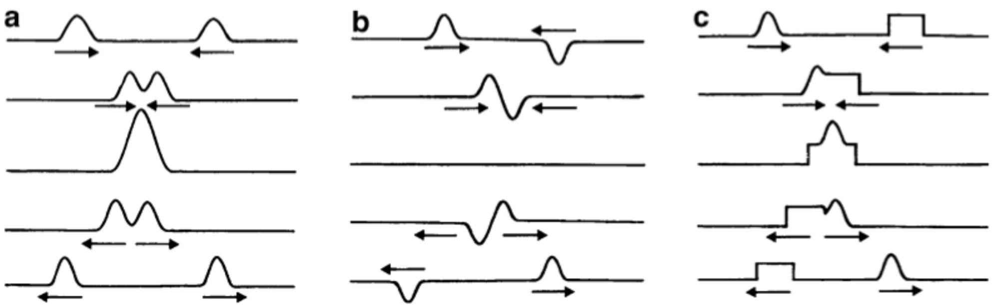

Figure 2.1: The superposition of travelling waves: (a) pulses in same phase; (b) pulses in opposite phase; (c) pulses with different shape. Adopted from [1]

The string can vibrate in three different polarization planes: vertical, horizontal, and longitudinal. Vertical and horizontal waves are called transverse waves, as their vibration planes are perpendicular to the direction of propagation. The waves propagating lengthwise along the string are called longitudinal waves [21].

The longitudinal waves play an important role in the sound of the attack and create the metallic character of the low notes [22]. The longitudinal component is generated from the transverse vibration by the nonlinearity of the string [23]. In case of the piano, the hammer excites one polarization of the string, and the other two gain their energy through coupling [22]. The longitudinal waves appear at the sum and difference frequencies of the two transverse waves [23].

2.2 End Conditions

When the propagating waves reach the ends of the string they reflect back [1]. The shape of the reflected wave depends on the end conditions of the string. A fixed end means that the string termination is securely fastened. When the wave reaches the fixed end of the string it reflects back to the sender. At the fixed end the displacement y is zero, and the general solution is

$$y = 0 = f_1(ct - 0) + f_2(ct + 0), \tag{2.3}$$

from which

$$f_1(ct) = -f_2(ct). \tag{2.4}$$

Thus, the reflected pulse at the fixed end is turned upside-down.

In contrast to the fixed end, at the free end the displacement can move freely but the tension needs to be maintained in the x direction. Because no transverse force is possible the spatial derivative $\delta y/\delta x = 0$. Thus,

$$y = f_1'(ct) = f_2'(ct). \tag{2.5}$$

Integration gives

$$f_1(ct) = f_2(ct). \tag{2.6}$$

Thus, a pulse reflects without a sign change on a free end. Figure 2.2 shows how the pulse is reflected at the fixed end and at the free end.

Free and fixed end conditions are only the special cases. On many stringed instruments such as the kantele the end points are only nearly fixed as they bend slightly [19]. However, for simplicity, strings of plucked stringed instruments are often treated as fixed from both ends.

2.3 Standing Waves

In a string that is fixed from both ends the travelling waves reflect as they reach the termination points of the string. Interference of the waves travelling to opposite directions leads to standing waves [18]. In standing waves there are points where the waves always sum up and on other points where they always cancel each other. The point in which the displacement is always zero is called a node. Between the nodes there are anti nodes in which the amplitude reaches its maximum value. The maximum amplitude is the sum

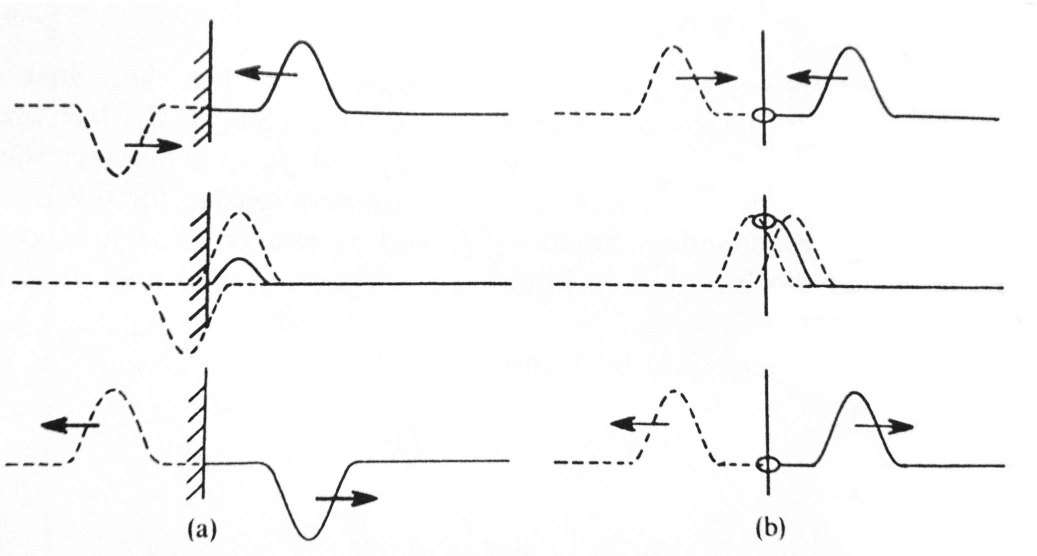

Figure 2.2: Reflection of a pulse: (a) fixed end; (b) free end. Adopted from [19].

of the individual impulse amplitudes and the oscillation frequency is the same as the frequency of the original impulses.

2.4 Plucked String

When a string is plucked the resulting vibration can be considered as a combination of multiple modes of vibration [18]. The mode with the lowest frequency is called the fundamental frequency and in an ideal case the frequencies of the rest of the modes are whole-number multiples of the fundamental frequency. These higher modes are called harmonics, the fundamental frequency being the first harmonic. The term partial includes also frequencies that are not whole-number multiples.

The resulting shape of a vibrating string can be attained as the sum of every vibrating mode at the time instant [19]. However, stringed instruments can be excited in a way that particular modes have zero amplitude [7]. The plucking point affects the modes in a way that if the string is plucked at one-third of its length, every third harmonic is missing from the resulting sound. Similarly, if the string is plucked at one-fifth of the string, every fifth harmonic is missing. Musicians can take advantage of this feature, as they can alter the plucking point based on the sound they want to produce. Plucking at the center creates a bell-like tone whereas plucking from the end of the string creates a sharper and fuller tone, as there are more modes present. Figure 2.3 shows the modes of a string that is plucked at its center. There are only odd harmonics with altering phase. Additionally, the relative amplitudes of the modes decrease as the order increases.

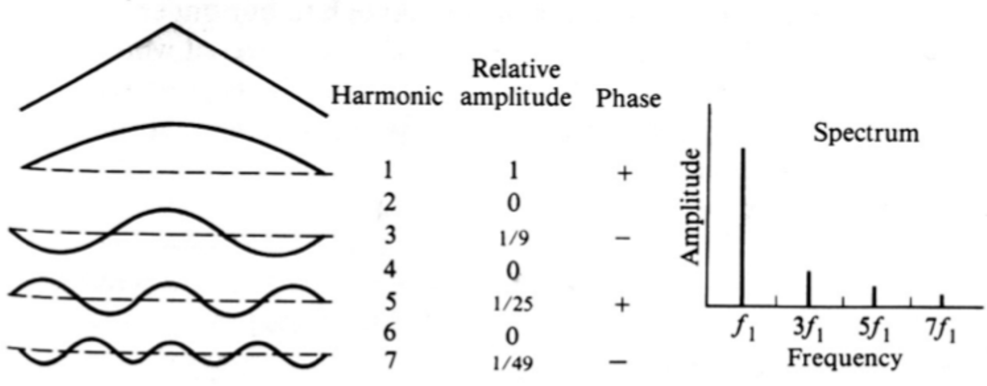

Figure 2.3: Harmonics of a string plucked at its center. Due to the plucking point the even harmonics are missing. Adopted from [19].

2.5 String-Obstacle Collisions

Collisions have a significant role in many musical instruments [15]. When an otherwise ideal freely vibrating string collides with an obstacle, such as the frets, a curved bridge, or the distorters studied in this thesis, that obstacle acts as a new termination point [10]. The obstacle limits the amplitude of the vibrating string creating nonlinearities. However, the string touches the obstacle only a short period of time and meanwhile the string continues its free vibrating motion [10]. Bilbao [15] studied the collisions with the string and the fretboard. Figure 2.4 shows how the shape of the string is altered as it collides with the string. With higher plucking force the effect is more significant.

The timbres of Indian stringed instruments, such as the biwa or the shamisen, are strictly determined by their string termination [14]. Their curved bridge acts as a nonlinear termination point as it interacts with the high amplitude waves and alters the shape of the string. The length of the string alters as it attaches to and detaches from the curved bridge [24]. Figure 2.5 shows stroboscopic images of the vibrating string without and with a curved bridge.

As the amplitude of the vibration decreases, the amount of high-frequency components increases [14]. Lower frequency components have often higher

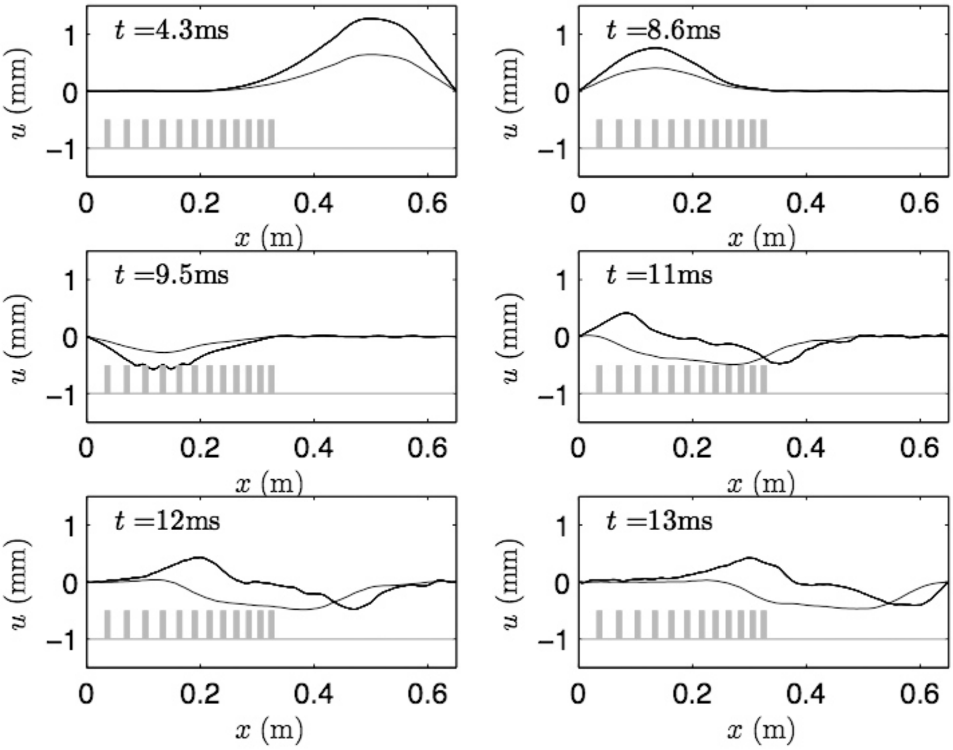

Figure 2.4: The evolution of the profile of a string in contact with the fretboard. The string is plucked with maximal excitation forces of 0.5 N (thin line) and 1 N (thick line). Adopted from [15].

vibration amplitude than high frequency components, which causes them to collide with the mechanical distorter [25]. As the amplitude of a wave decreases, the energy is transferred to higher partials, because high frequency waves have already passed the point of impact and do not attenuate on collision [14]. This is caused by the dispersion, which means that the transfer speed of a wave depends on its frequency: The high-frequency waves are faster than low-frequency waves.

2.6 Damping Mechanisms

The string vibrations lose their energy via damping mechanisms. Three main damping mechanisms are internal damping, air damping and energy loss through supports [26]. Internal damping consists of viscoelastic and thermoelastic losses. A string is strained as it is stretched and returns to

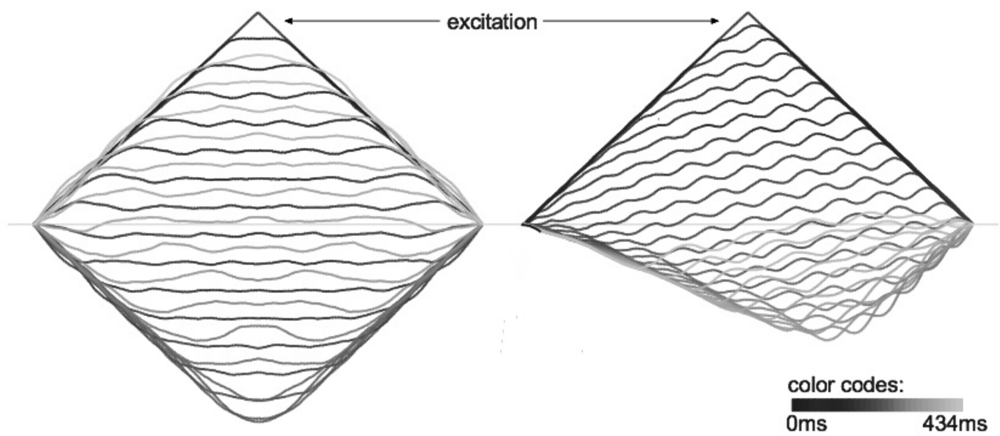

Figure 2.5: Stroboscopic images of the string without a curved bridge (left) and with a curved bridge (right). Adapted from [25].

its original state as the stress is removed. Viscoelastic losses are caused by a small delay between strain and stress. Thermoelastic losses are caused by the lack of thermal equilibrium between the compressed hot regions and extended cool regions. Viscoelastic losses affect the high frequencies whereas thermoelastic losses affect the middle frequencies.

Air damping includes the air absorption and viscous damping. Since the strings are poor radiators, the air absorption has hardly any effect. The air around the vibrating string causes viscous damping as it resists the movement. The energy losses through supports are determined by the mechanical characteristics of the bridge and the body. In addition, nonlinearities of the system introduce additional damping [26].

2.7 Body

The vibrating string itself radiates very little sound so the vibration has to be amplified with the body of the instrument. On some stringed instruments, such as the violin and the guitar, which is shown in Fig. 2.6, the other end of the string lies on the bridge. The bridge transmits the energy of the vibrating string to the top plate, which is often the main radiator at the high frequencies [1]. Simultaneously the top plate transmits energy to the back plate and the air cavity [18].

The top plate, the side plates and the back plate form the air cavity. The sound radiates from the plates as well as through the sound hole. The lowest mode of the air cavity is the Helmholtz resonance, which is determined by the diameter of the sound hole and the volume of the air cavity [18]. At low frequencies the bridge is perceived as a part of the top plate and most of the sound is radiated from the plates and the sound hole [26]. Each mode has an individual radiation pattern based on the shape of the body. Figure 2.7 shows how radiation patterns of a guitar depend on the frequency.

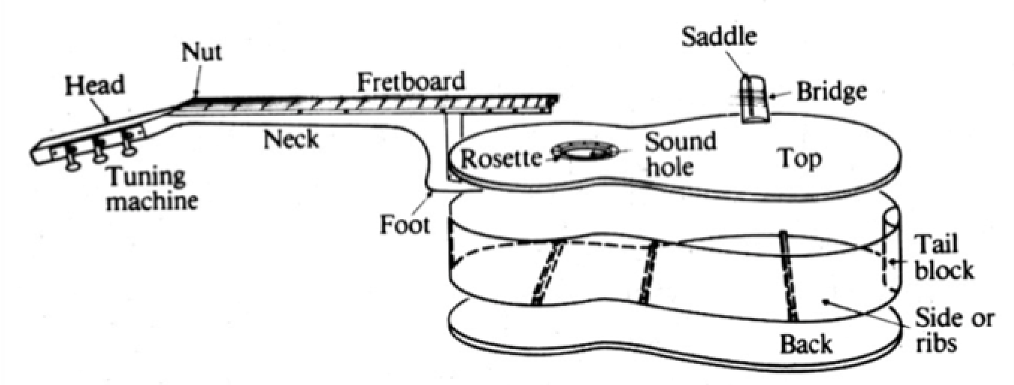

Figure 2.6: An expanded view of a guitar. Adopted from [1].

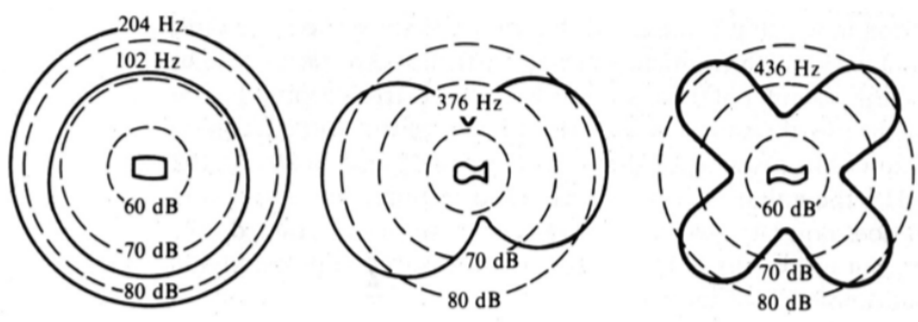

Figure 2.7: Sound radiation patterns of a guitar on different frequencies. Adopted from [1].

Some instruments, such as the kantele from the zither family, have their strings attached straight to the body. However, many stringed instruments have a neck that includes the tuning pegs and the nut. Some of the instruments, such as the guitar and the sitar, have also frets on the neck, which simplifies the playing [18]. They are used to split the string in order to produce tones with different pitches. One problem of the frets is that the pressing of the string against the fret increases tension of the string. This increases the pitch slightly making the note flat. However, addition of extra length in between the frets usually compensates this.

Chapter 3

Hearing

Sound can be defined as vibrations of 20 to 20 0000 Hertz that travel through an elastic material [18]. Even though this can be easily measured, it does not explain how human actually senses the sound. However, sound can also be defined as the sensation that vibrations stimulate in the organs of hearing [27]. This refers to the perception inside of the ear and is hard to quantify. The study that investigates the relationship between the physical stimuli and the sensations and perceptions is called psychophysics.

Human hearing has been developed to receive sound waves that are limited between 20 Hz and 20 000 Hz. The amplitude is limited by the hearing threshold and pain threshold and the dynamic range is expressed in decibels [28]. The hearing threshold is the minimum sound level a human can hear. The most intense sound human can hear without damaging an ear, the pain threshold, is about 120 decibels above the hearing threshold, which means that the human hearing has a wide dynamic range [29].

By having two ears, humans can locate the direction of the source based on the level differences and the phase differences between the ears. At high frequencies above 4000 Hz the head masks the sound source from the other ear and the localization is based on the loudness difference. At low frequencies below the 1000 Hz the localization is based on the phase difference or the arrival time of the sound. Between 1000 and 4000 Hz, the localization has the worst accuracy. Additionally, a pair of ears makes the hearing system more reliable since human can hear even if one ear is damaged. Besides localization, the human auditory system is selective [18]. The listener can hear a single sound from large amount of sounds, for example the voice of the speaker in a noisy environment such as a cocktail party.

3.1 The Anatomy of the Ear

Purpose of the human ear is to receive audio waves that are travelling in the air. The ear is divided into three parts: the outer ear, the middle ear, and the inner ear. Figure 3.1 shows a schematic diagram of the ear. The middle ear and the inner ear have been enlarged for the purposes of illustration.

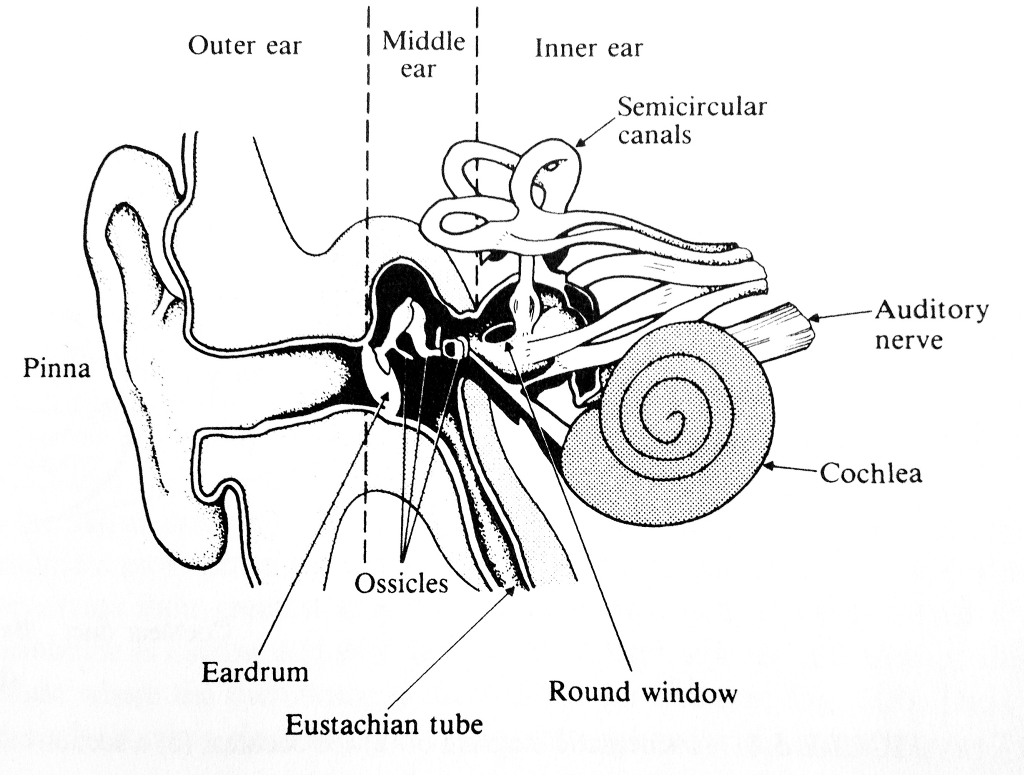

Figure 3.1: A schematic diagram of the ear. Adopted from [18].

3.1.1 Outer Ear

The outer ear consists of pinna and auditory canal, which are fully linear systems [18, 28]. The outer ear is a passive system that terminates to the eardrum. The external pinna collects sounds and enables us to locate the direction of the sound by masking the high frequencies. Without the pinna it would be impossible to distinguish if the sound arrives from 45 degrees or from 135 degrees angle. Humans are very sensitive to the changes in pinna and even small changes complicate the detection of the direction [28].

The ear canal can be treated as a hard walled acoustic tube. The ear canal has a resonance at frequency of about 4000 Hz and it boosts the hearing sensitivity in the range from about 2000 to 5000 Hz. The boost can be up to 10 decibels. However, the ear canal has no effect on the location detection. Transfer functions, the hearing sensitivities at different frequencies, of human ears vary between humans. Around 1000-2000 Hz the differences are only few decibels while on the range from 3000 to 15 kHz differences can be from 20 to 30 decibels [28].

3.1.2 Middle Ear

The middle ear starts from the eardrum that transfers pressure variations of the sound waves in to mechanical vibrations [18]. The eardrum is attached to the three small bones called ossicles: malleus, incus and stapes. Their purpose is to perform impedance adaptation between the outer ear and the inner ear. The middle ear is connected to the inner ear via stapes that touches the oval window of the cochlea. Since the area of the stapes is much smaller than the area of the eardrum, the pressure on the oval window is enhanced.

In addition to the impedance adaptation the ossicles protect the inner ear from extremely loud noises. So called stapedius reflex stiffens the three bones thus attenuating the incoming sound [28]. As the eardrum tightens the stapes is pulled away from the oval window. However, the reflex has an effect only at the lower frequencies and operates quite slow so it does not fully prevent from hearing injuries.

In addition, the middle ear contains Eustachian tube that connects the middle ear to the upper throat. Its purpose is to equalize the pressure between the outer ear and the middle ear. This is necessary, since the pressure difference shifts the eardrum from its original position, which increases the impedance. High impedance decreases the sensitivity of the hearing especially at the lower frequencies. Compared to the original state, the tones at frequency of 500 Hz are attenuated from 12 to 15 decibels and tones at frequency 1500 Hz are attenuated from 0 to 6 decibels [28]. Intensities at frequencies over 2 kHz remain unchanged.

3.1.3 Inner Ear

The inner ear consists of the semicircular canals and the cochlea [18, 28]. The semicircular canals are necessary for balance but have no effect on hearing. The purpose of the cochlea is to transfer the pressure variations into neural impulses. The cochlea is a fluid filled spiral with 2.7 rounds and with a length of 35 mm. It is connected to the middle ear by the oval and the round window. The oval window is covered by the stapes while an elastic membrane covers the round window.

The cochlea, shown in Fig. 3.2, is divided into two chambers by the basilar membrane which vibrates from the movement of the fluid. Different frequencies vibrate different parts of the membrane. The high frequency tones create highest amplitude near the oval window where the membrane is narrow whereas the low frequency tones vibrate the other end where the membrane is thick. The basilar membrane rests on the organ of Corti, which contains about 20 000 to 30 000 small hair cells that are evenly spread. As the basilar membrane vibrates it creates horizontal movement which bends the hair cells. This activity sends impulses to the brain via the auditory nerve. The impulses include data of the location of the vibration, the style and the amplitude. Auditory nerves transfer the information to the central auditory system via the cortex.

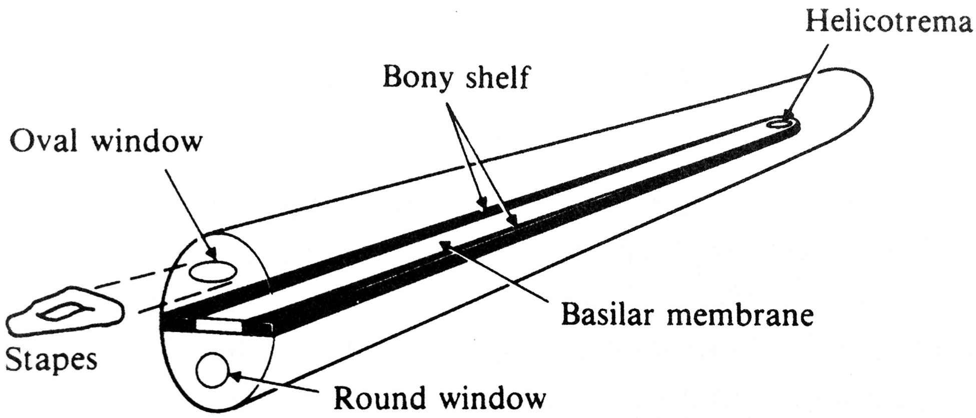

Figure 3.2: A schematic of uncoiled cochlea. Adopted from [18].

3.2 Sound Power Level, Sound Intensity Level and Sound Pressure Level

Sound power level describes the power of the sound source in decibels [18]. Even though decibels are often used to compare two quantities, they can be used as an absolute value with a reference value. Sound power levels are expressed in decibels by using $W_0 = 10^{-12}$ W as a reference value. The equation for sound power level is

$$L_W = 10 \log \frac{W}{W_0}, \tag{3.1}$$

in which L_W is the sound power level in decibels, W is the power of the measured sound, and W_0 is the reference value.

Intensity describes the sound at a specific point. It can be expressed by comparing the measured intensity to the reference intensity $I_0 = 10^{-12}$ W/m^2. Thus, sound intensity is defined as

$$L_I = 10 \log \frac{I}{I_0}, \tag{3.2}$$

in which L_I is the sound intensity level, I is the measured intensity, and I_0 is the reference intensity. Intensity depends on both the surrounding sound field and the sound pressure of the source.

Human ear reacts to the air pressure changes in a sound wave. The intensity of a sound wave is proportional to pressure squared. The equation for intensity of a sound wave is defined as

$$I = p^2/\rho c_0, \tag{3.3}$$

where I is the intensity, p is the sound pressure, ρ is the density of air, and c_0 is the speed of sound in air. They both depend on the temperature and at normal temperatures the product is from 410 to 420. In calculations the product value 400 is used for ρc_0. Sound pressure level (SPL) is defined as

$$L_p = 10 \log \frac{p^2}{400 I_0} = 10 \log \frac{p^2}{4 \times 10^{-10}} = 20 \log \frac{p}{p_0}, \tag{3.4}$$

where the reference level $p_0 = 2 \times 10^{-5}$ N / m^2 = 20 µPa. The reference level 20 µPa on 1 kHz is also the hearing threshold, the minimum sound level human can hear. It is selected as the zero value of the decibel scale in sound pressure level measurements [28].

3.3 Perception of Loudness

Loudness is an attribute that is used when sounds are ordered from quiet to loud [29]. It is a subjective quantity and high sound intensity level or sound pressure level do not necessarily mean that the sound is loud. Besides the intensity, loudness depends also on the frequency content and the quality of the sound. Loudness depends also on the duration of the sound, it grows up to 200 ms [18]. After that the duration has no effect on the perception.

3.3.1 Equal-Loudness Curves

Equal-loudness curves in Fig. 3.3 show how the hearing sensitivity depends on the frequency. The curves are built with hearing tests in which the subject has to match the test tone with an equally loud 1000 Hz tone [29]. The level of the 1000 Hz tone defines the loudness level of the test sound and it is measured in phons. Loudness scale was developed to be able to quantity loudness of tones. Its unit is sone which is defined as 40 dB 1000 Hz tone which equals to 40 phons.

The thresholds are plotted rising upwards and the lowest curve is the absolute threshold curve describing the lowest sound level a human can hear. Thresholds increase rapidly at very high frequencies and at low frequencies. The human hearing is most sensitive at frequencies in between 2000 Hz and 5000 kHz. It reaches its maximum between 3000 Hz and 4000 Hz which is near the first resonance frequency of the outer ear canal [18]. The sensitivity peaks also at about 13 kHz which is the second resonance of the ear canal. There is a slight dip on the 500 Hz. Basically, low tones need more intensity to sound as loud as middle tones. Based on the equal-loudness curves the instrument is perceived loud if it has lots of energy at frequencies between 2000 Hz and 5000 Hz. Since low frequency components need lots of energy to sound loud, mechanical distorters should be used to transfer their energy to the higher frequencies. This way the same energy would produce a louder sound. Age affects on the highest tones that a human can hear. Young children can hear tones up to 20 kHz whereas the high end of most adults is limited in between of 10 kHz and 15 kHz. [29]

3.3.2 Sound Weighting

Loudness of a sound can also be measured with a sound level meter that measures the sound pressure level. Since the hearing sensitivity varies with frequency, to get a truthful presentation of the loudness, summing of the intensities of the frequencies is not enough. To overcome this, different weightings with predefined coefficients have been introduced [29]. Standardized A, B, C, and D weighting networks all approximate hearing sensitivity with slightly different criteria. A, B and C networks are shown in Fig. 3.4. The most used weighting is A-weighting that is based on the 30 phon equal-loudness contour. It is used in loud environments such as construction sites and loud concerts, in which the loudness level must not be exceeded. However, it is only a rough estimate of the human hearing as it attenuates low and very high frequencies and gives a slight boost on high middle frequencies.

Sound level meters using different weightings produce a reliable results

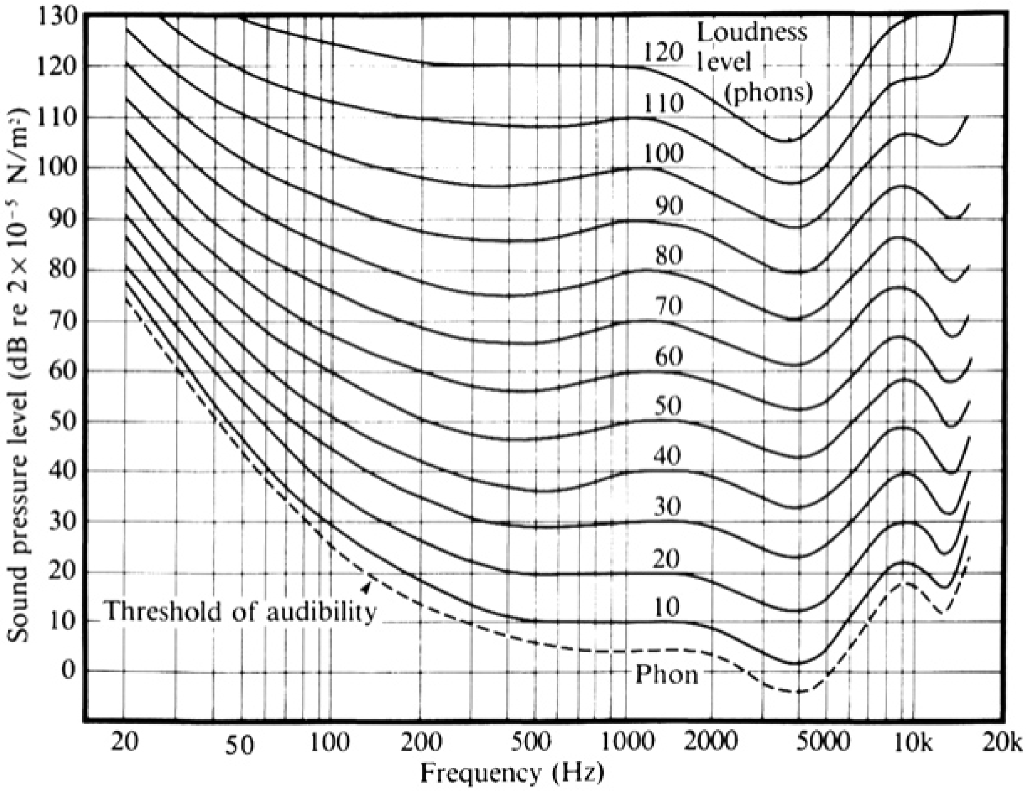

Figure 3.3: Equal-loudness curves. Adopted from [18].

only with relatively long duration sounds [29]. They cannot produce a good estimate of the loudness of transient sounds or when the sounds are divided over widely separated frequency bands [29].

3.3.3 Critical Band

If multiple frequency peaks are introduced close to an existing peak, the perceived sound level does not change. However, as the combined sound becomes more broadband the sound appears to be louder. Bandwidth at which the signal loudness ceases to increase is called the critical band [28, 29]. The critical band describes the resolution of hearing and defines how wide frequency range human perceives as one with a constant loudness. The width of the critical band is close to constant at low frequencies. It expands as the frequency rises being about a one-third of an octave, and the audible range comprises about 24 critical bands. However, locations of critical bands are not constant as they are dynamically formed around the frequency peaks. In general, human hearing processes sounds in a way that every sound inside a

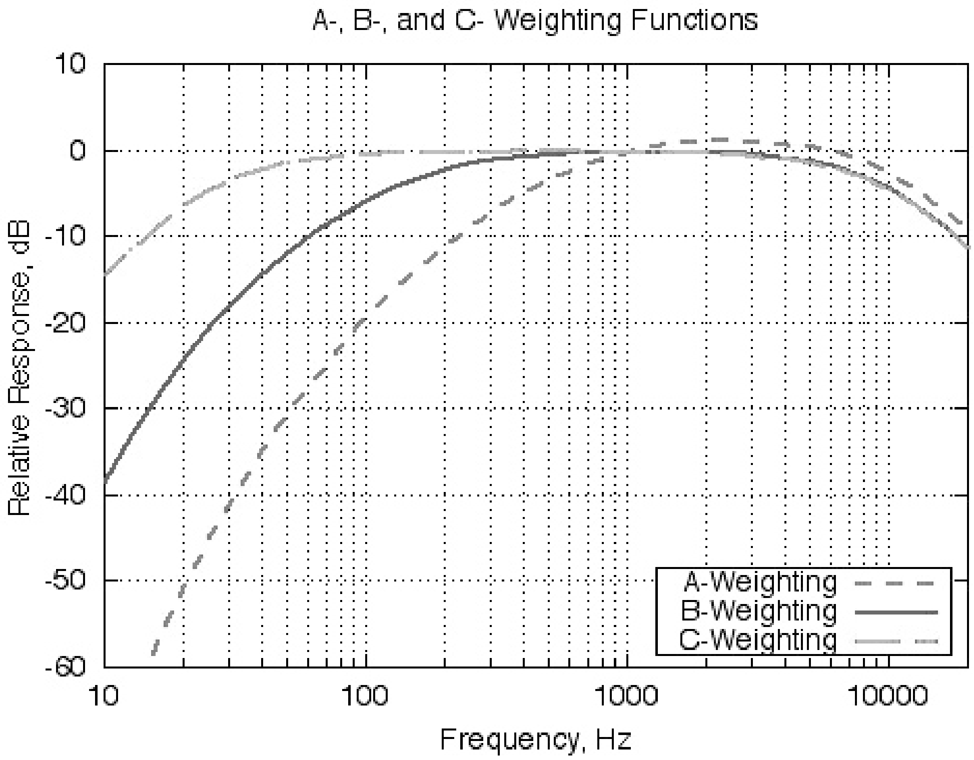

Figure 3.4: A, B and C weighting networks. Adopted from [30].

critical band is considered as one sound.

Broadband sounds seem to be louder than narrowband noise even though they have the same sound pressure level [18]. Due to this a musical instrument should produce as broadband sound as possible in order to be as loud as possible. The more critical bands the sound passes, the louder it is perceived. If there are two individual tones that are very far from each other, only the louder one defines the overall loudness [18].

3.3.4 Masking

In noisy environments quiet sounds cannot be heard under the background noise. This is due to an effect called masking. Masking means that the threshold of audibility of a sound is raised when another masking tone is present [18, 29]. Masking can be caused by a simultaneous sound. Sound is most easily masked by noise which contain similar frequency components. If the noise and the sound are widely separated in frequency, no masking occurs. Tones with higher frequency than the masking noise are masked more effectively than tones with lower frequency. White broadband noise shows an approximately linear effect on the threshold level, as the noise level increases the threshold level increases the same amount. In addition to increasing the threshold level, masking reduces the loudness.

Additionally, the masking can occur in two directions in time: forward and backward masking. In forward masking the sound is masked by an earlier sound that ends 20 to 30 ms before the sound begins. This suggests that hair cells need time to recover after stimuli. In backward masking the signal is masked by a masking signal that ignites after the desired signal. The amount of masking decreases as the time interval between the sounds increases.

3.4 Timbre

Timbre is used to denote the tone quality or the color of a sound [18]. Timbre is defined as that attribute of auditory sensation of which a listener can judge that two similarly presented sounds that have the same loudness and pitch are dissimilar [29]. Timbre makes it possible to distinguish between the same note of different instruments and instrument families. The distorters studied in this thesis alter the sound of the instrument.

A single tone can be described with frequency and intensity, but all of the natural tones are more complex. Timbre cannot be ordered based on one attribute such as loudness or pitch, instead it is multidimensional [27]. Timbre depends on the spectrum of stimulus, waveform, sound pressure, spread of frequencies and temporal characteristics of the stimulus. Timbre is related to the relative level produced by a sound in each critical band and the dimensions to characterize timbre is limited by the number of critical bands [29].

The instrument tones can be divided in to three segments: the attack, the steady-state, and the decay [31]. A tone starts with the attack, which includes the initial transient, which is a high amplitude short-duration sound. Next is the steady-state, in which the stability is reached at the end of the attack. The decay defines how the tone trails off. The attack segment is the most important for the identification of a tone.

Timbres of steady tones are determined by their magnitude spectra. Besides the constant spectrum, timbre depends on transients and fluctuations that vary over time [18]. Recognition of musical instruments depends on whether the sound is periodic or more noise-like. It depends also on the transient spectrum and the structure of the envelope. For example a piano tone has a rapid onset and a gradual delay. Many instruments have noise-like qualities such as the flute that has a small puff preceding every tone [29]. If the transient is removed from a sound the identification of the instrument becomes difficult.

The most common way to order timbres is the usage of verbal attributes [18]. Dull-sharp scale has been found to be the most significant. Three other often used attributes are brightness, fullness and roughness [31]. Brightness is related to the midpoint of the energy distribution. Fullness depends on the relative precence of odd and even harmonics. Tones that consist only of odd harmonics sound hollow or nasal if many harmonics are present. Roughness is present in tones which have strong high harmonics, above the sixth harmonic. High harmonics make the tone sound sharp and penetrating. Tones with strong low harmonics, which are below the sixth harmonic, sound mellow [29]. They sound richer and more musical than simple tones still remaining soft, since the higher harmonics are absent. The strong fundamental mode makes the tone sound rich.

Chapter 4

Distortion in Musical Instruments

Distortion refers to modification of an audio signal [32]. Typical audio systems try to minimize the amount of distortion and in the most of the fields distortion is characterized as an unwanted change to the signal [33]. However, there are situations in which distortion is a desired feature, and probably the most well known is the distorted electric guitar. Distortion creates a sound that is full of high frequency partials, and it is used to produce more powerful or harsh tones [34].

In audio the most common source of distortion is clipping, in which the maximum amplitude of a wave is limited. Figure 4.1 shows how clipping affects the waveform of the audio. As the input signal exceeds the maximum value, the nonlinear distortion processes can produce output signals that have energy on the frequencies that had no energy in the input [34]. The clipping is divided into hard clipping and soft clipping. In hard clipping there is a sharp discontinuity at the clipping point, which produces harsh sounding high amplitude odd harmonics. On contrary, in soft clipping the transition is smoother, which reduces the amount of odd harmonics [35].

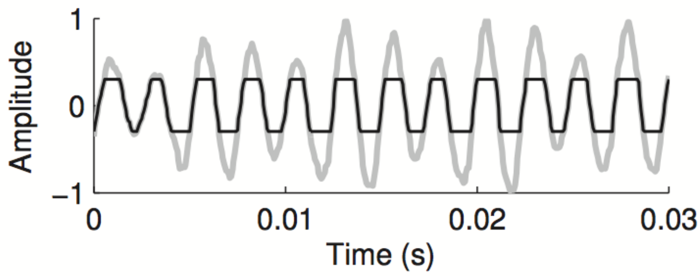

Figure 4.1: A signal (gray) and its clipped version (black). Adopted from [36].

When the additional frequencies are a whole number multiples of the input frequencies, the distortion is called harmonic distortion [32]. Figure 4.2 shows the frequency domain plots of the undistorted and distorted signals. Those frequencies are musically related to the input signal and affect the timbre of the instrument. Nonlinearities in audio systems are often measured in terms of harmonics added to the pure sine wave fed to the system. Harmonic distortion may be expressed as the relative strength of individual components, in decibels, or the root mean square of all harmonic components [37].

In intermodulation distortion at least two input signal frequencies form new components that are possibly inharmonic [32]. Instead of being purely harmonic, the new frequencies are also sum and difference of the original frequencies. In Western music the inharmonic components tend to sound off-key.

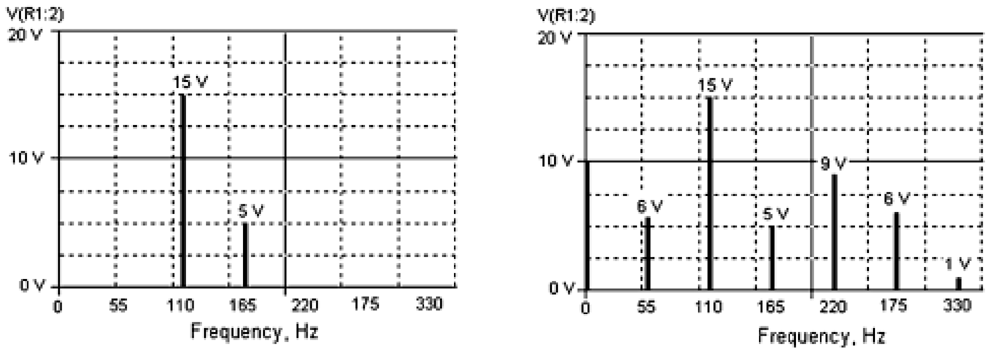

Figure 4.2: Frequency domain plots of the undistorted (left) and distorted (right) signals. Adopted from [35].

4.1 Electrical Distortion

Electrical audio distortion refers to any kind of deformation of the output compared to the input such as clipping or harmonic distortion caused by nonlinear behavior of the electrical components [37]. Clipping occurs as the audio signal is overdriven and amplified beyond the maximum capacity of the system. The system amplifies the signal only to its maximum value and cuts off the rest of the signal, which forces a sine wave to a distorted square wave type waveform. The resulting signal has additional high frequency partials compared to the original signal.

The first distorted guitar sound was produced by a broken speaker cone [33]. The producer Sam Phillips liked the sound so much that he wanted to use the sound. The next milestone was when amplifiers were purposely altered to produce distorted tones. The guitarists found the distortion interesting. At first the distortion was introduced as powerful guitar pickups were routed through less powerful amplifiers. There are different types of distortion for electric guitar. Overdrive is a distortion that has more compression, distortion and sustain in the sound, but maintains the character of the original sound. An effect that is so distorted that the character of the sound is changed is referred as distortion [33].

Vacuum tube amplifiers are widely used in professional guitar amplifiers [38]. Their soft clipping produces pleasant even-order harmonic distorted tones that are preferred over the odd-order harmonic distorted tones of the solid-state electronics. Guitarists often use multiple gain stages to distort their tone. Besides the amplifier, guitar effect pedals such as overdrives, distorters and fuzz pedals can be used to produce distortion. Most of the distortion applications work based on an electrical trigger, which is often Schmitt's trigger [39]. The trigger creates a square wave output which is summed with the signal of the guitar.

The direct distorted guitar signal sounds harsh and unpleasant. However, the guitar loudspeaker acts as a low pass filter with a cut off frequency of about 5000 Hz. This attenuates most of the high partials and leaves a pleasant thick and musical sound. Additionally, the distortion pedals are built to filter out the highest frequencies in order to produce a smoother timbre [33].

4.2 Mechanical Distortion

Even though the electrical distortion is what comes to mind when the word distortion is mentioned, the first broken speaker cone used for distortion was actually a mechanical distorter. A mechanical distorter is a device that creates nonlinear behavior in an acoustic instrument. Raman [40] noted that the curved bridge of the tanpura, an Indian stringed instrument, created powerful harmonics which are absent in the case of an ideal string. As the string collides with the curved surface the vibration amplitude decreases and the string starts to vibrate at a more rapid rate. In this procedure the energy of the fundamental is transferred into energy of the harmonics. Addition of the harmonics makes the sound of the instrument brighter and broadband, thus louder for human hearing.

The nonlinear artefacts are similar to the artefacts generated by amplitude clipping in analog to digital conversion [41]. Similarly as in electrical distorters, the string gets square-wave like properties that create distortion by adding harmonics that were not part of the original signal. The nature of distortion depends on the plucking energy and the duration of the vibration [41]. As the vibration amplitude increases, the interaction with the string and the obstacle is increased as well. The vibration amplitude decreases due to damping effects, which reduces the amount of collisions.

In the Western stringed instruments the goal has always been to create as clean timbre as possible. Strings can vibrate freely and there exists no nonlinearities. On the other hand, Indian instruments, for example, are known for their buzzing timbre which give more character to their sound.

There are many reasons why the timbres are so different between Western and Eastern instruments. One reason is of course the musical taste for such sounds. Another suggestion states that it is because the Greek amphitheaters were built to emphasize higher frequencies [1]. Because the wind and chattering lies on the lower frequencies, the stairs row was built to work as a high pass filter. As a result, the low frequencies are reduced and higher frequencies are enhanced. However, the Eastern instruments such as the Japanese biwa and the Indian sitar were usually played on the streets. The instruments had to be loud in order no to be masked by the environment noise. This could be achieved by creating a bright and broadband timbre that cuts through the noise.

Many kinds of different mechanical distorters exist and they affect the timbre of the instrument significantly. The following instruments with nonlinear features are presented: The biwa, the shamisen, and the sitar are Indian instruments that have a recognizable buzzing tone. They all have a curved bridge that creates the high harmonics. The traditional Northern Europe instrument the kantele, which is a member of the zither family, has unusual termination points that create nonlinearities [16]. The slap bass has a strong nonlinear interaction with the strings and the fretboard. The last is the widely used Western instrument the piano and especially the part-pedaling technique, in which the string collides with an obstacle.

4.2.1 Biwa

The biwa is a Japanese stringed instrument which is plucked with a plectrum [42]. It consists of four strings and four frets and its length is less than 1 meter [43]. Figure 4.3 shows a photo of a biwa. The tuning depends on the players voice, for example A3 (220 Hz), E3 (165 Hz), A3 (220 Hz) and E4 (330 Hz). The frets of biwa are relatively high and far from each other. The body and the bridge of biwa is constructed to support frequencies over

500 Hz. The strings do not produce enough high frequency components and therefore a sawari mechanism is used to create a buzzing sound.

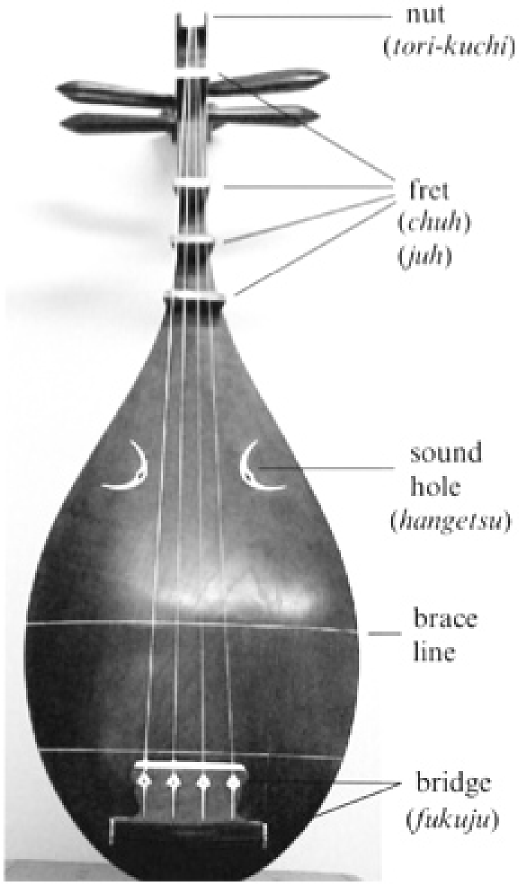

Figure 4.3: Biwa. Adopted from [44].

Sawari means an obstacle or touch and it can be interpreted as a purposely way to create more character to the timbre [43]. The sawari is created by slightly curving the surface of the nut and frets. As the string repeatedly collides with the surface, the surface generates few corners to the string on every vibration period and excites nonlinear high-frequency components. The sawari intensifies higher partials and generates rich spectral components and also prolongs their duration. The amount of extension of decay time and higher harmonics depends mainly on curvature of the surface and secondly on the thickness of the string. The wider the strip of the sawari surface is, the stronger the sawari effect. The sawari is applied to all biwa tones, four frets and the nut.

While the total energy of the string vibration remains constant, the maximum vibration amplitude decreases. According to Taguti the sawari intensifies partials from 6th to 20th partial, even to 24th partial and the spectral structure changes over time [12, 42]. However, no sub harmonics appear due to sawari [12].

4.2.2 Shamisen

The shamisen, shown in Fig. 4.4, is a Japanese plucked stringed instrument that has a long unfretted wooden neck and a small body [43]. The strings are tuned from about 140 Hz to 400 Hz and they are plucked with a large plectrum. The front and back are covered with white cat skin and the bridge is placed on the skin. The striking vibrates the cat skin membrane which makes part of the characteristic sound of the shamisen.

Figure 4.4: Tsugaru Shamisen. Adopted from [45].

The shamisen has a rich sound with strongly emphasized high harmonics caused by the sawari effect [46]. The sawari effect was first found in the biwa and was then reproduced in the shamisen. Whereas the biwa had the sawari on each string on nut and frets, the shamisen has the sawari only on the lowest string. The lowest string bypasses the nut and touches and separates from it as the string vibrates. The sawari effect is affected by the height of the lowest string.

The sound of shamisen resembles the banjo but the sawari mechanism makes it sound also like the sitar. The averaged spectrum of shamisen is triangularly shaped with the top located at 700 Hz, which corresponds to the strongest resonance of the whole instrument [43]. The sawari effect strongly emphasizes the partials above the eight harmonic. The shamisen introduces also inharmonic noise-like components at frequencies above 2 kHz. They are caused due the complex reaction with the bridge and the membrane surface as well as due to the elliptic string motion [43]. The elliptic string motion exists because the waves in Z-polarization have the different soundboard impedance than the waves in X-polarization.

4.2.3 Sitar

The sitar shown in Fig. 4.5 is a plucked lute used mainly in Hindustani and Indian music. It has a resonating chamber and a long hollow fretted neck. Number of the frets depends on the type of the sitar and they are all movable. The sitar can have 18 to 20 metallic strings which are plucked with a metallic plectrum. It has either 7 main strings and 3 drone strings or 3 main strings and 4 drone strings [47]. Additionally, there are 11 to 13 sympathetic strings that run underneath the frets.

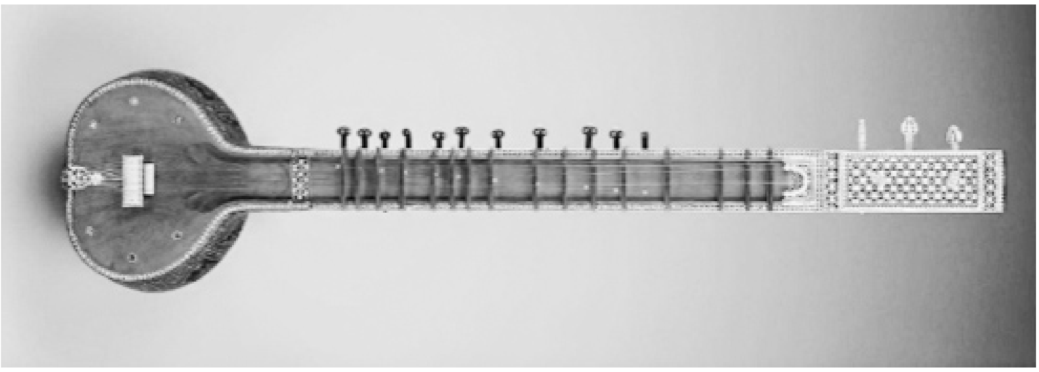

Figure 4.5: Sitar. Adopted from [44].

The sitar has two characteristic features: The first is the sympathetic strings which produce an echo like effect as the energy of the main strings transfers to them. The second is the curved bridge called the jawari which produces the spectrally rich buzzing tone of which the sitar is famous for. When the string vibrates its length alters as it touches the curved bridge, which produces long-lasting high harmonics [41]. The frequencies of these harmonics are dynamic, since the clipping occurs at different levels depending on the point where the string touches the bridge [41]. The adjustment of the curved bridge is very important for the musicians and they are in a close collaboration with the maker of the sitar [24]. The sympathetic strings have also a curved bridge called the taraf that introduces the buzzing tone on the decaying part of the sound. Figure 4.6 shows the jawari and the taraf of the sitar [48].

The sound of the sitar is bright nasal and its spectral centroid is relatively high [47]. Plucking of a sitar string at a certain node does not mute the corresponding node, which is the case with an ideal string. Additionally, the

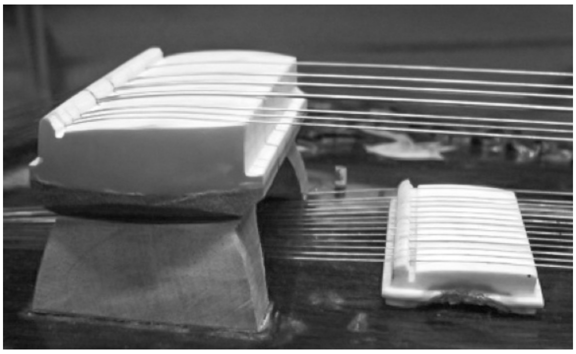

Figure 4.6: Jawari, the curved bridge for the main strings of the sitar, and taraf, the curved bridge for the sympathetic strings. Adopted from [48].

amplitude clipping occurs at different lengths of the string as the contact point between the string and the bridge varies [41]. For that reason the frequencies of the harmonics descend over time, which can be easily heard and can be seen from the spectrogram in Fig. 4.7 [25].

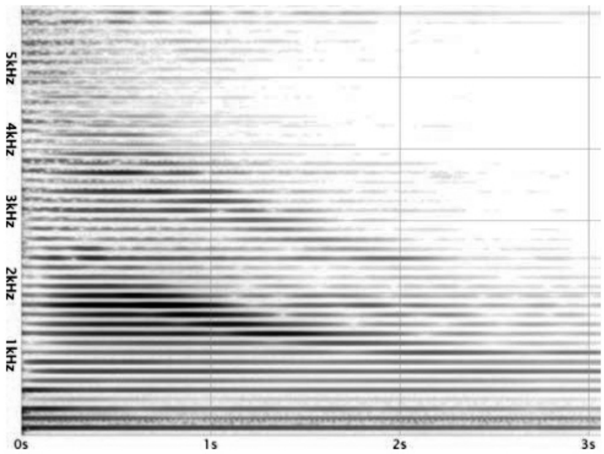

Figure 4.7: Spectrogram shows the descending formants of the sitar. Adopted from [25].

There are also other sitar-like Indian instruments such as the Indian drone lute tanpura, the sarangi, the rudra vina, and the sarasvati vina [47]. The tanpura has a similar curved bridge as the sitar, but it features also a thread placed in between of the string and the bridge [49]. The thread is typically made of silk or cotton. Van Waltsijn and Chatziioannou [49] show that the combination of the bridge and the thread create a precursor of strong high formants with fast decay. The desired amount of jawari can be adjusted by altering the place of the thread. The sarangi and the two vinas are equipped with either sympathetic string jawari or main string jawari, but not both.

4.2.4 Kantele

The kantele shown in Fig. 4.8 is a traditional stringed instrument from the Northern Europe with strings that span the length of the sound box and sound board. A kantele belongs to the family of zither instruments and it has 5 to 40 strings and a wooden body [16, 50]. The historic kantele was built out of one piece of carved wood, which was open from the bottom, and had strings made out of horsehair. The modernized version of the kantele has metal strings and is built of two or more pieces of wood. The bottom of the kantele may be open or closed, in which case there is an air cavity inside. The kantele is played with either fingers of left and right hand or strummed with a piece of wood or a plectrum.

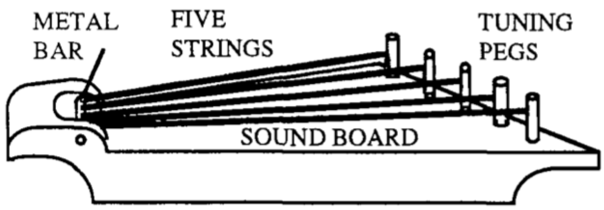

Figure 4.8: Structure of a 5-string kantele. Adopted from [16].

The kantele has an unusual way the strings are fastened and supported [16]. The strings are fixed around a metal bar called varras by a knot and wounded around a tuning peg at the other end. These terminations are not rigid end supports and they have influence on the effective length of the string as it vibrates. The knotted termination produces a strong beating, which enhances the second harmonic component [51]. The effective termination points of the string are different in the two main vibration planes. For vertical polarization the knot shown in Fig. 4.9 is the termination point, whereas for the horizontal polarization it is the metal bar.

Sympathetic free strings are characteristics for the kantele and they produce the brilliant reverberant timbre. The reverberation is mostly caused by the common metal bar that connects multiple strings together and transfers the energy of a the vibrating string to the other strings [16]. The energy transfer of a 5-string kantele is the strongest between the fifth and the first string, followed by the fifth and the third, the fourth and the second, and the third and the first string [51].

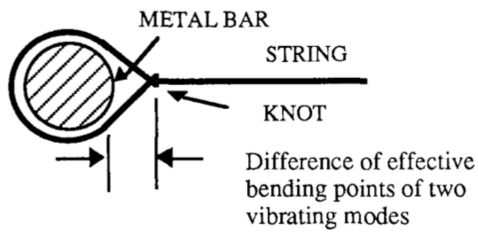

Figure 4.9: String with a know termination around the metal bar of kantele. Adopted from [16].

4.2.5 Slap Bass

The construction of the bass guitar is similar to the guitar, but it has only four low-tuned strings and a longer neck. The bass guitar is usually plucked with fingers or a plectrum.

Bass slapping is a playing technique which produces a brilliant and percussive sound [9]. The string is either struck against the body with the thumb or it is plucked heavily with the index or the middle finger. The slapping causes the string to collide with the fretboard or the frets, which leads to nonlinear limitation of the signal. Multiple collisions can occur simultaneously [11]. The collisions are close to inelastic on fretless instruments and elastic on fretted instruments. On fretless fingerboard the sound is less bright since wider area of the string collides with the fingerboard.

The frets and the fingerboard react as another reflection point as the string collides with them. Amplitude limitation adds broadband noise pulses to the signal. The signal evolves during the first few periods and becomes more regular as the amplitude decreases. The high frequencies decay quickly after the excitation.

4.2.6 Piano

The piano is an instrument that is played using a keyboard. As a key is pressed down the hammer inside the piano strikes the metal string. The vibration transfers to the wooden body which radiates the amplified sound. The sound attenuates when the key is lifted up. However, the note can be sustained by pressing the sustain pedal at the bottom of the piano.

Piano has some nonlinearities. After hitting the string the hammer stays in touch with the string and reacts as another reflection point [9]. The effect is similar as in the slap bass. In contrary to the slap bass frets, the hammer creates only a temporary reflection point. Also the other end of the treble strings of the piano is bent over V-shaped capo bar. Slight curvature on the capo bar makes the string act nonlinearly and the spectral components evolve over time [14]. The curved bridge adds high-frequency oscillations to the tone similarly to the curved bridge of the sitar. The amount of additional spectral components depends on the force of the hammer blow [14].

Additionally, the sustain pedal can be used to create multiple artistic expressions [13]. Instead of depressing and releasing the pedal, it can be used for vibration or part-pedaling. In part-pedaling the sustain pedal is pressed slightly against the strings which introduces nonlinearities to the signal. Especially in the bass strings the nonlinear amplitude limitation transfers energy from low frequencies to the higher partials, which can excite the missing modes. Figure 4.10 shows how part-pedaling ignites the partial that is located at the striking point and would be absent with full pedal. However, with part-pedaling the partials decay faster than with full pedal.

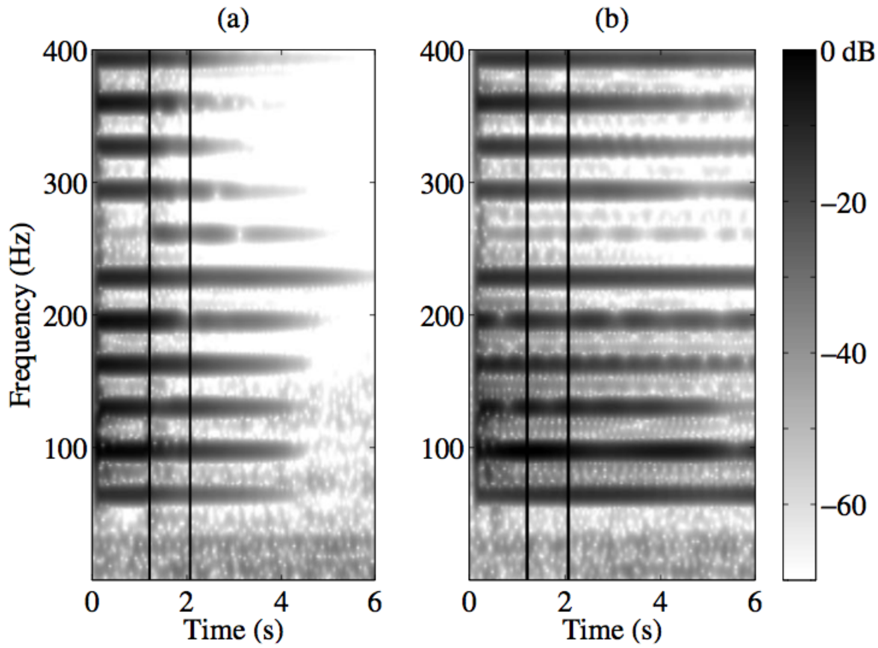

Figure 4.10: Spectrograms of tone C1: (a) with part-pedaling; (b) with full pedal. Spectrograms show that part-pedaling ignites the eight harmonic which is absent with full pedal. Vertical lines show the duration of damper-string interaction with part-pedaling. The same interval is shown in (b) for reference. Adopted from [13].

Chapter 5

Measurements

5.1 The Kantele

The kantele studied in this thesis is built by Jyrki Pölkki and is shown in Fig. 5.1. It is a traditional kantele that has a solid top plate without the sound hole and a carved bottom. The kantele has eleven strings which are tuned from A3 to D5. The strings are made of metal and they are stretched between two separate metal bars and the tuning pegs. The fifth string was used in these measurements. It is 420 mm long and was tuned to 269.2 Hz, which is a sharp C4 (261.63 Hz).

Figure 5.1: The Kantele

5.2 The Measured Distorters

Three distorters were measured. They all interact with the string in different ways and were made of different materials. In all distorters, the string comes in contact with the distorter multiples times as it vibrates. Figure 5.2 presents the distorters D1, D2 and D3 accordingly and shows how they were in contact with the string. The distance between the distorters and the tuning pin was 20 mm to 30 mm.

Distorter D1 is a curved metal wire, a paper clip, touches the string only at one point. The touching point is at 20 mm distance from the tuning pin. Placement of D1 is quite difficult and it is very sensitive to changes. Only a small nudge from proper placement could either mute the string completely or leave the string undistorted. Additionally, too strong plucking force could displace the distorter from its place.

Distorter D2 is also made out of curved metal wire, but it has two touching points with the string. The touching points were located at 20 mm and 25 mm distance from the tuning pin. As the string vibrates it interacts with both of them. Placement of the D2 was easier than with D1 and the distorter settled on quickly. With D2 it was easier to achieve a similar sound if the distorter was taken off and put back. At the latest when the string is plucked few times the D2 adapts to its place.

Distorter D3 is a piece of rawhide that was attached to a metal wire. The surface that touches the string is 10 mm wide and it is spread from 20 mm to 30 mm from the tuning pin. D3 produces a sitar-like sound. Similarly to D1 and D2, placement of D3 has a strong effect on the distortion. If the distorter touches the string too much it mutes the whole string. Slight changes produce various timbres and the user has to decide the optimal one.

5.3 The Measurement Setup

In order to minimize the reflections of the room, the kantele was measured in a small anechoic chamber that is anechoic above 125 Hz. The kantele was placed on a chair near the wall. The measurements were done with a Rode NT1-A condenser microphone that was placed one meter above the chair. The recording setup is shown in Fig. 5.3. Rode NT1-A has a low self-noise level and a relatively flat frequency response [52]. The microphone was connected to Motu Ultralite MK3 audio interface.

The audio samples were recorded with Apple Logic Pro X digital audio workstation with the sample rate of 44.1 kHz. The recordings were done in 24 bits to ensure sufficient quality. The gain of the microphone amplifier was

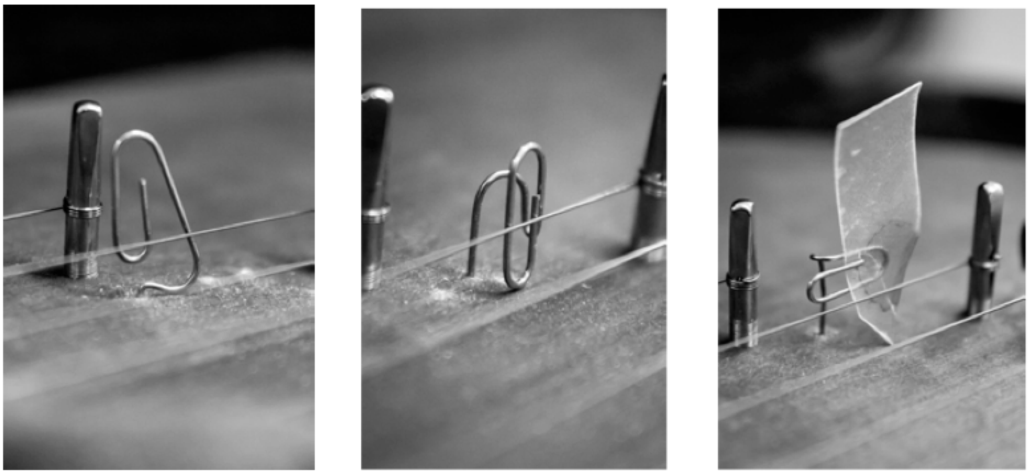

Figure 5.2: The distorters, from left to right: D1, D2 and D3.

adjusted once and then kept the same so that all the measurements would be comparable with each other. The A-weighted level of undistorted string of the kantele was measured with Brüel & Kjær Type 2250 sound level meter with Type 4189 microphone. The A-weighted level measured at one meter above the kantele peaked at about 70 dB. The gain was adjusted so that signal would peak at -10 decibels.

All the measurements were done on the fifth string to ensure that each measurement would be easily comparable. Each distorter and plucking point combination was measured at least three times to ensure that there is at least one good sample of each distorter and plucking point. Even though sympathetic strings are a characteristic feature of the kantele, the rest of the strings were muted with foam since only effects on one string was studied. Strings were muted at many points to ensure that the foam does not cause any accidental natural harmonics.

Copper wire was used to give required impulse to the string. Copper wire breaks at a repeatable level of stress so it is suitable for scientific measurements [53]. Since 0.15 mm copper wire was too durable and gave too strong impulse, 0.1 mm copper wire was used. The 0.1-mm wire gave optimal force of about 3 newtons being loud enough but it did not make the string touch the body of the kantele. Quality of the distortion depends on the used plucking force, so it would have been good to measure with altering forces. However, there were no suitable copper wires in between of 0.1 mm and 0.15 mm.

Often stringed instruments are measured with impulses coming straight from above. However, distorters were set in such way that they react only

Figure 5.3: The measurement setup of the kantele in an anechoic chamber.

to horizontal vibrations. If impulses were given right from the top there would have not been any distortion. As a result the impulses were given in a constant angle which was approximately 45 degrees. This is close to the actual plucking direction of the kantele. Four fingers on top of each other were used as a reference point to ensure the proper angle on each measurement.

The kantele was measured with two different plucking points: one at the center of the string and another at one-third of the string. The plucking points were selected in such manner that they attenuate as many partials as possible. At first, the plucking point was at the center of the string, which in ideal case attenuates all of the even harmonics, ending up with a bell like tone. The second plucking point was at the one-third of the string, which attenuates every third harmonic partial in an ideal case. It is also a common plucking point by the kantele players.

Chapter 6

Results

The measurements are studied by using four different methods, and the results were partially presented in [54]. The first method is the spectrogram which visually represents how the frequency content of the sound varies over time. The features studied were the harmonic partials. The magnitude and its evolution over a time frame of each individual partial is presented visually and the time of decay is calculated. The third feature is the temporal evolution of the spectral centroid, which is a good measurement for the brightness of a sound [55]. The last method is the comparison of the A-weighted sound pressure levels. A-weighting is used to approximate the perceived loudness of the sound.

Typically mean of multiple samples is studied, since it reduces the effect of divergent plucking. However, since distortion has a large variation in between individual plucks, only one sample of each distorter was used for the results. This provides a sufficient amount of information for comparison between the undistorted string and the other distorters.

6.1 Spectrogram

A spectrogram presents the amount of energy in a frequency band as a function of time [29]. The intensity of the signal is expressed in grey scale in which black denotes the highest intensity and white the lowest intensity. The spectrogram is highly usable for distortion measurements, since it shows how the energy is divided over harmonic partials and shows their duration in a single figure.

The measured signals were filtered with a fourth-order Butterworth high-pass filter with cut-off frequency at 150 Hz in order to remove the entire unnecessary low frequency hum. Since the fundamental frequency of the string is 269.2 Hz this has no affect on the spectrum of the string. The spectrograms of the signals were calculated with a window size of 600 samples and with an overlap of 570 samples. Figure 6.1 shows the resulting spectrograms of the string plucked at the center. The lowest intensity values were limited to -120 dB in order to achieve a clear figure. The difference between the black and white areas is 60 decibels. The frequency scales of the figures are limited to frequencies under 12 000 Hz. The frequencies above that are not so crucial for the human hearing. In addition, the figure is clearer when the band is narrower.

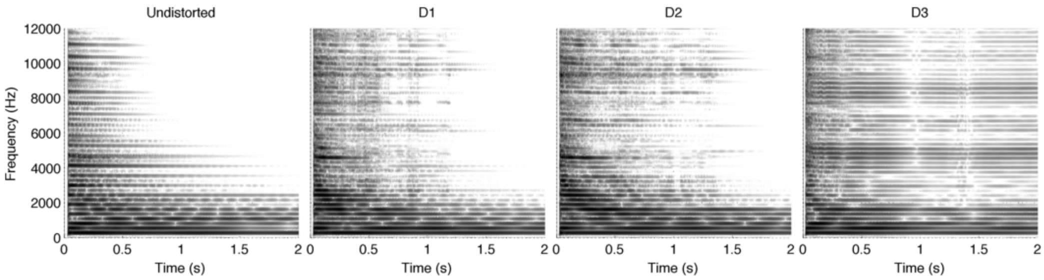

Figure 6.1: Spectrograms of string plucked at the center of the string

The spectrogram of the undistorted string shows that the most of the energy is located at the low end of the frequency spectrum. Higher harmonic partials decay faster than lower partials. The partials below 2.5 kHz maintain their energy for over 2 seconds. From 2.5 kHz to 5 kHz the harmonic partials decay after 1.5 seconds. Up till 10 kHz harmonic partials have energy for over half a second and the partials above 10 kHz attenuate after 0.5 seconds or faster. There are strong high-frequency partials at frequencies 4121 Hz, which is the 15th partial, 8360 Hz (31st), 9760 (37th), 10 370 Hz (39th) and 11 050 Hz (41st). All of the strong partials are odd partials.

According to the spectrograms, all of the distorters transfer energy from the low end to the high end of the spectrum. Similarly to the undistorted string, the strongest components are at the low end of the spectrum at the frequencies below 3000 Hz. However, there is noticeably more energy at high partials and they maintain their energy from one second to over two seconds. The higher partials of the D1 have the fastest decay time as they decay soon after one second. D2 decays marginally slower. However, D3 maintains it high-frequency energy for the whole two-second time window.

The shape of spectrograms of the distorters D1 and D2 are very similar. Both of them have some noticeable stronger partials and weaker partials. Both of them have a strong 17th partial which decays after half of a second. The spectrograms indicate that D1 and D2 have their strongest effects on partials above 8000 Hz. Especially harmonics in between 9500 Hz and 11

000 Hz last longer than most of the partials. Instead, the spectrogram of D3 has a different shape. There is no noticeable strong frequency peaks as there were on the spectrograms of D1 and D2. The partials between 4000 Hz and 5000 Hz and between 7800 Hz to 9000 Hz have the most energy of the higher frequencies. D3 ignites also the even high frequency partials. There are some partials that decay fast, for example 22th (5884 Hz), 27th (7232 Hz) and 37th (9876 Hz) partials.

All of the distorter spectrograms have noticeable gaps in the energy of the higher partials. D1 has two gaps, the first gap lasts from 0.5 seconds to 0.8 seconds and the second gap lasts from 0.9 seconds to 1.1 seconds. D2 does not have as clear gaps, but there can be seen one from 0.35 seconds to 0.45 seconds and another short one is after one second. D3 has very clear gaps, as energy of all of the high frequencies seem to be eliminated for short periods of time. The first gap is at 0.3 seconds, the second is at 1 second and the third one is at 1.45 seconds. The second and the third gap are more noticeable than the first one.

The gaps might occur as the vibrating string detaches from the distorter for short periods of time. Then the string can vibrate freely since the distorter does not limit the amplitude. When the string touches the string again the higher partials are back due to the nonlinear behavior of the string. When the vibration amplitude decreases the distorters no longer limit the vibration and the high partials disappear.

Figure 6.2 illustrates the spectrograms of the string as it was plucked at one-third of its length. The spectrogram of the undistorted signal shows that every third partial is slightly attenuated. However, since kantele has nonlinearities, they are not completely muted. The effects of the distorters are slightly different compared to the pluck at the center but the trend is the same. Each of the distorters enhances the high frequency partials and expands their decay time.

Similarly to the spectrogram of the string plucked at the center, most of the energy is at the low frequencies below 3000 Hz. Up till 7000 Hz partials maintain their energy from one to 1.5 seconds. The partials on frequencies above 7000 Hz decay before 0.75 seconds. There are few stronger high frequency partials: the 10th, the 13th, the 31st and the 34th. Partials close to 8000 Hz are not as strong as they were when the string was plucked at the center.

D1 has the most long-lasting energy at the higher partials and the gaps cease to exist. The reason for this could be that the distorter touches the string better than when the string was plucked at the center. The whole spectrum range is utilized and there is no more clear strong peaks or attenuated partials. The sixth partial is strong even though it is attenuated on

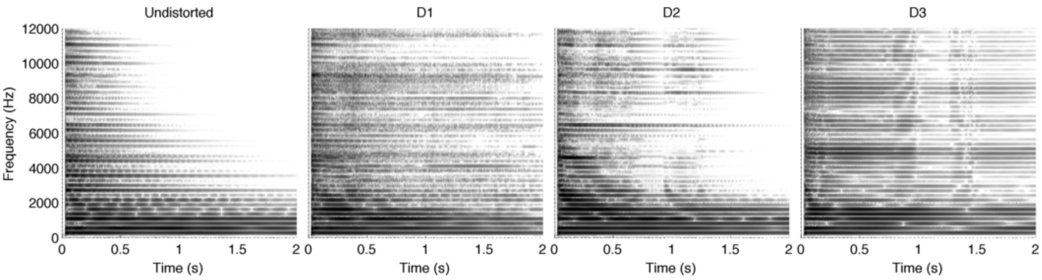

Figure 6.2: Spectrograms of string plucked at one-third of the string

undistorted string.

The spectrogram of D2 is quite similar than before when the string was plucked at the center. D2 excites high-frequency partials with long-lasting energy and there are some gaps in the spectrogram. When the string is plucked at one-third of its length D2 lacks some high-end energy but has a stronger 23th partial than before when the string was plucked at the center. The rest of the strongest partials are the same as in the earlier case: 29th, 33rd, and 37th.

D3 acts similarly on both plucking points. However, at the time of 0.9 seconds pitch of the partials ascend on the frequencies above 2000 Hz. From one second to 1.3 the partials maintain their high pitch until at the time of 1.3 seconds the pitches of the partials begin to descend back to their original states. This might be caused by the wide surface of rawhide. As the string attaches to it, the surface gradually shortens the effective length of the string. This effect is unique to D3 and cannot be reproduced by distorters D1 or D2. Spectrograms show that the gaps on the string plucked at the center and the ascending and descending of the string when the string is plucked at one-third of its length occur at the same time. On those time periods the string detaches from the string.

As a conclusion, according to the spectrograms all of the distorters transfer energy from the low frequencies to the high frequencies. Since the human ear is more sensitive on higher frequencies, the sound may be perceived louder. The distorters create also a more broadband sound, which is perceived louder as the distorted sound exceeds more critical bands than the undistorted sound.

6.2 Harmonic Partials

The plots of isolated harmonic partials show how each of them evolve over time. The time is located on the x-axis and magnitude on the y-axis. The individual harmonics of the signal were extracted using a high-order fractional delay filter that cancels the neighboring harmonics [56]. The extraction shows the intensity and the decay time of the harmonics.

Time to decay 60 decibels, the T60 time, was measured with the T30 time. The T30 estimates the T60 time from time to decay 30 decibels, which is multiplied by two [18]. The measured T30 time of each partial is shown in the upper left corner of each plot. The amplitude maximum of each harmonic was used as the starting point and a first-degree curve was fitted between it and the point in which the amplitude was 30 decibels lower. The T60 time was calculated based on the achieved slope of the curve. The strong amplitude fluctuation caused errors on some of the distorted string T60 times as the attenuation of 30 decibels might have been reached too early.

Figure 6.3 illustrates the harmonic partials of the undistorted string. The harmonic partials reveal that most of the energy is on the low frequency partials, especially on partials below the 17th partial. The plots of the individual partials show that many of the even harmonics have almost as much energy as odd harmonics. However, above the 10th harmonic the trend is that even harmonics decay faster than the odd ones. The difference in decay time between the odd and the even harmonics is clearest on the harmonics between the 11th and the 16th harmonic. The decay rate of every partial is close to constant and there is strong ripple only on the lowest partials.

Figure 6.4 shows how energy is distributed over harmonic partials when the string is distorted with D1. The strongest components are located below the sixth harmonic similarly as on the undistorted string. However, the higher harmonic partials have more energy and decay slower than the high partials of the undistorted string. At partials below the 20th partial D1 gives similar results as the undistorted string. Instead of a constant rate, the decay rate varies but the overall trend is the same as with the undistorted string. However, partials above the 29th partial have noticeably more energy. Many of the high-frequency partials decay after 1.5 seconds which is one second longer than the corresponding decay time of the undistorted string.

D2 performs similarly as D1 and its harmonic partials are shown in Fig. 6.5. D2 distributes energy from the lower partials to the higher partials. Many of the harmonics decay slower than the partials of D1 and many of the upper harmonics maintain their energy for 1.5 seconds. The 20th partial of the D1 and D2 that decay faster than the rest.

The harmonic partials of D3 are shown in Fig. 6.6. D3 has the strongest high partials out of the three distorters. They decay slower and their level is higher than on the rest. There is a valley in the plot slightly before one second. In the spectrogram it was shown as a gap and these results show that there is less energy at high harmonic partials at that moment. The other gap

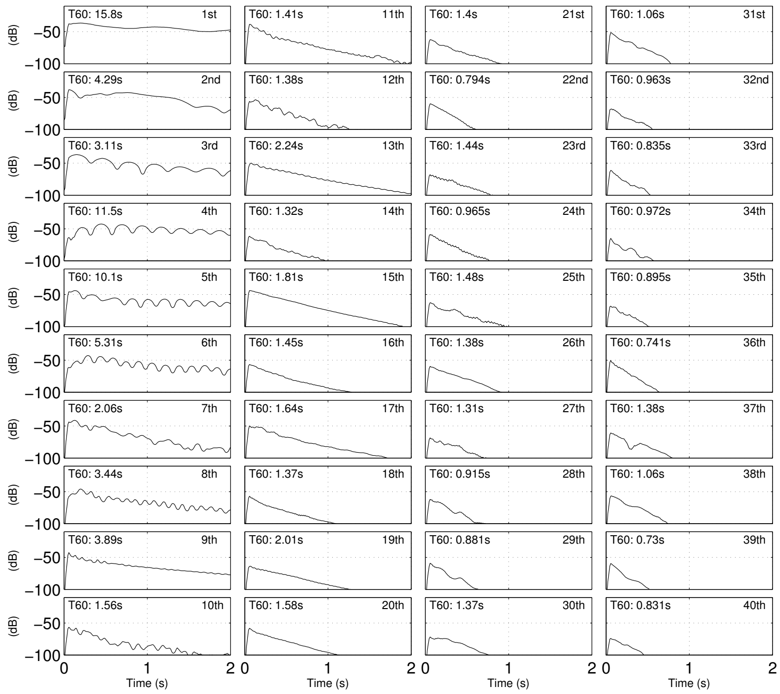

Figure 6.3: Harmonic partials of the undistorted string plucked at the center of the string

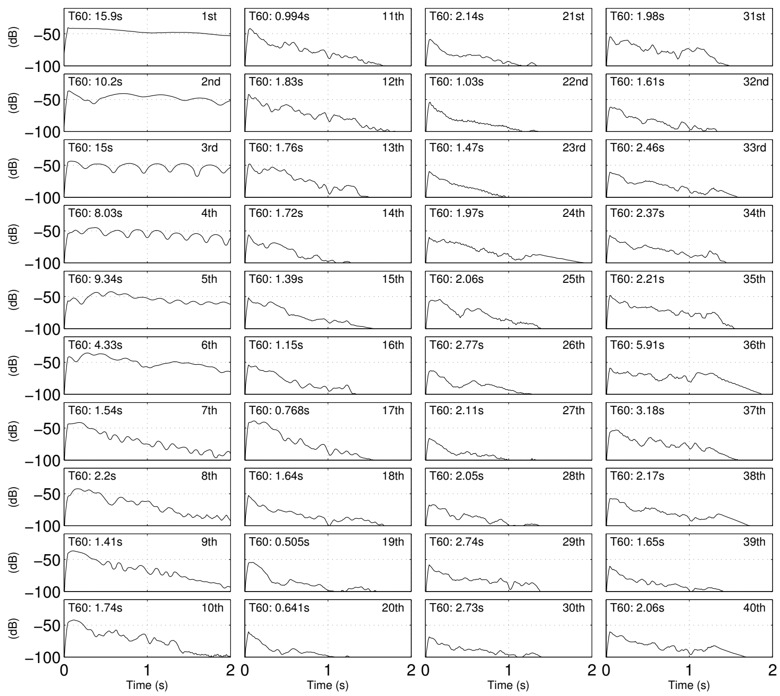

Figure 6.4: Harmonic partials of D1 plucked at the center of the string

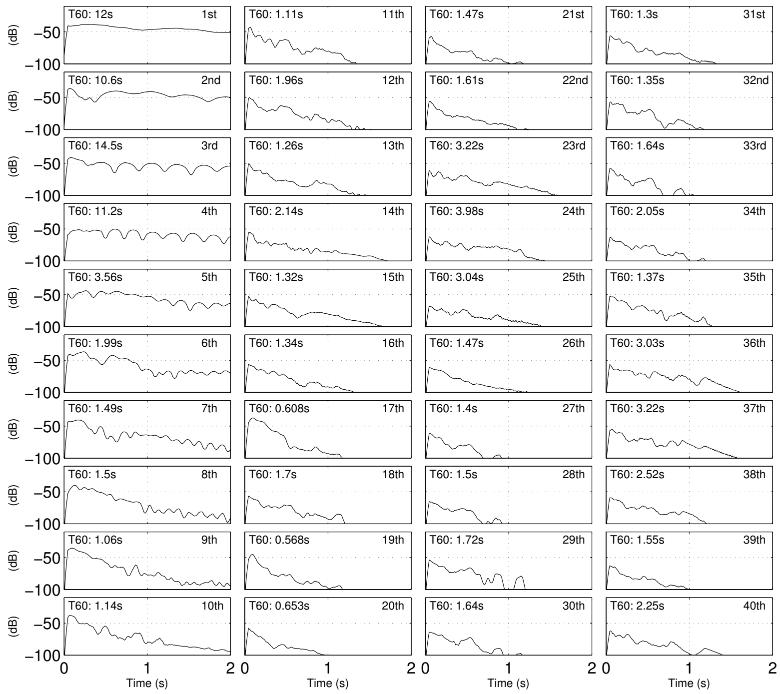

Figure 6.5: Harmonic partials of D2 plucked at the center of the string

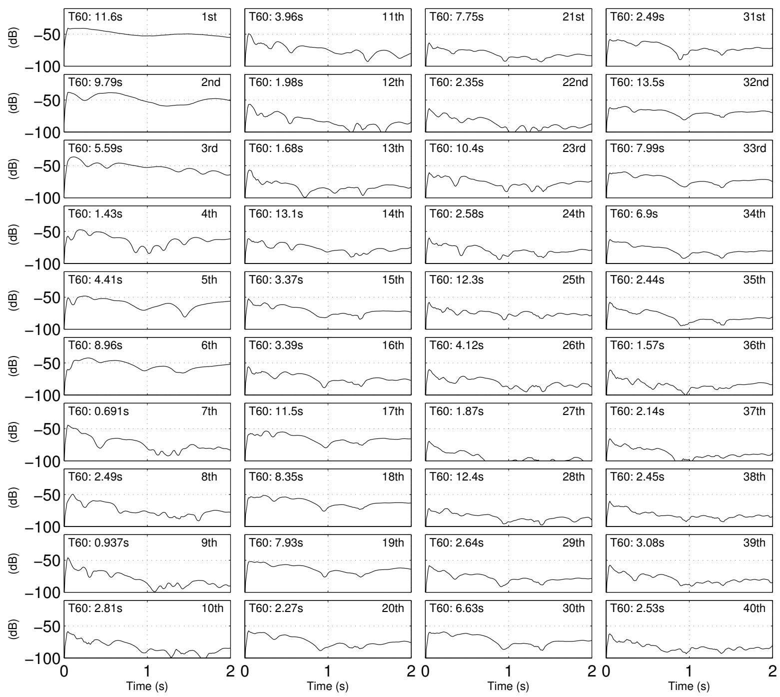

Figure 6.6: Harmonic partials of D3 plucked at the center of the string

at 1.5 seconds can also be seen as a valley in the magnitude. The gap effect can be seen clearly at the partials above the 15th partial. The 27th partial decays the fastest and the cause is not known.

Figure 6.7 shows that the undistorted string plucked at one-third of its length behaves similarly as the string plucked at the center. Some of the every third partial is attenuated, like 12th, 15th, and 18th. However, many of the partials seem to be unaffected by the plucking point because of the nonlinearities of the kantele. The partials start to decay faster above the 21st partial. However, there are some stronger harmonic partials on the high frequencies, such as the 23rd, the 26th, and the 37th.

The harmonic partials of D1 plucked at one-third of its length are shown in Fig. 6.8. D1 spreads the energy evenly on higher partials while the low partials below the sixth partial remain the strongest. However, there are no clear stronger or weaker partials. The partials decay at a constant rate even though their magnitude fluctuates. The T60 time of many partials is over two seconds, which is longer than when the string was plucked at the center.

D2 has more variation on the energy spread and its partials are shown in Fig. 6.9. There is lots of variation on the energies of the partials. There are some strong partials, such as the 24th, the 31st and the 36th, that decay slower and have higher energy levels than the other partials. The 20th and the 25th partial decay fast and have low energy. The gap, which could be seen in the spectrogram of D2, can be seen as an energy loss in harmonic partials. At the time before one second the energy of multiple partials decreases for a short period of time.

Figure 6.10 presents the harmonic partials of D3. The signal decays slowly compared to the other distorters or the undistorted string. The energy is divided quite equally over the whole frequency spectrum and there is no noticeable stronger or weaker harmonics. The ascending and descending of the pitch of the harmonic partials cannot be seen from the harmonic partials like it can be seen from the spectrogram. However, after one second the amount of energy decreases. This energy loss continues the whole time as the pitch remains high.

Harmonic partials show that the distorters transfer the energy to the higher partials thus increasing the perceived loudness. T60 times of high-frequency partials are increased from less than one second to over two seconds.

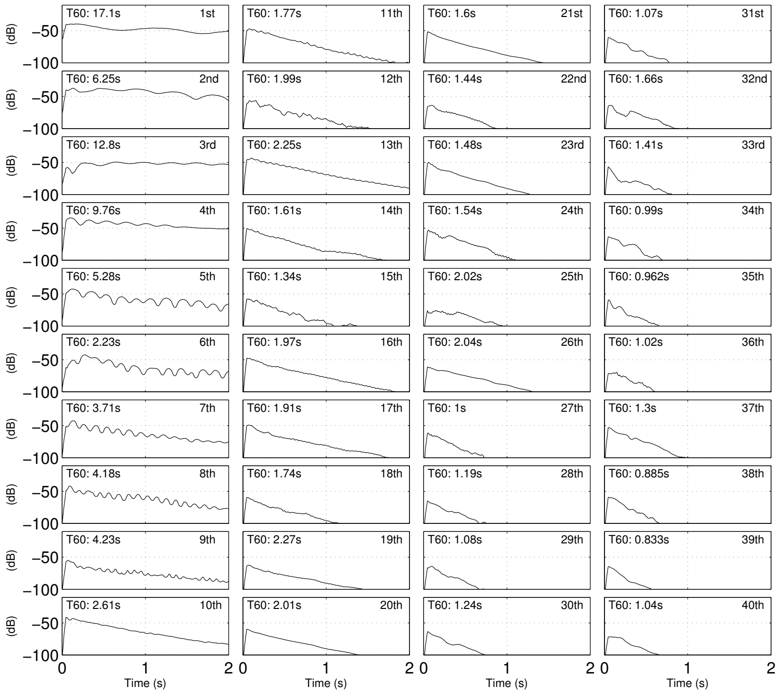

Figure 6.7: Harmonic partials of undistorted string plucked at one-third of the string

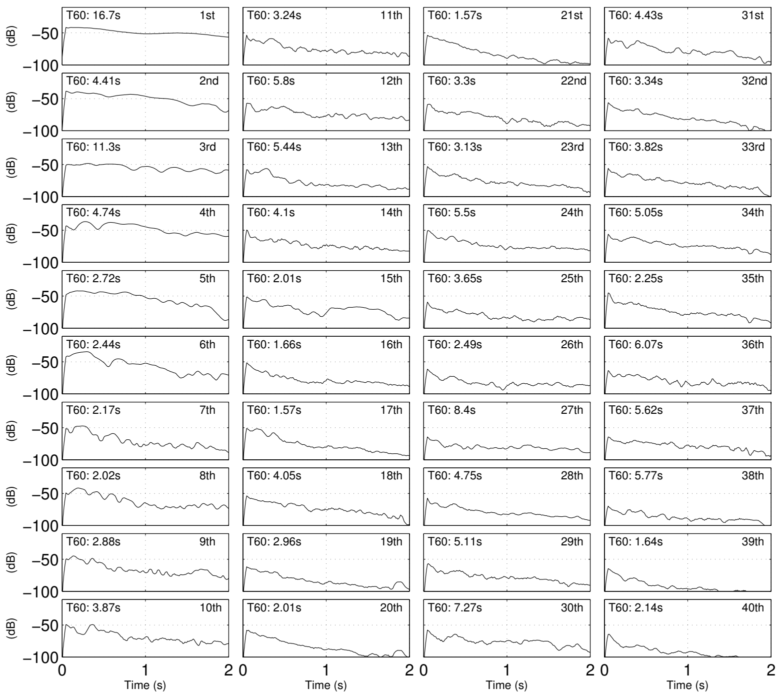

Figure 6.8: Harmonic partials of D1 plucked at one-third of the string

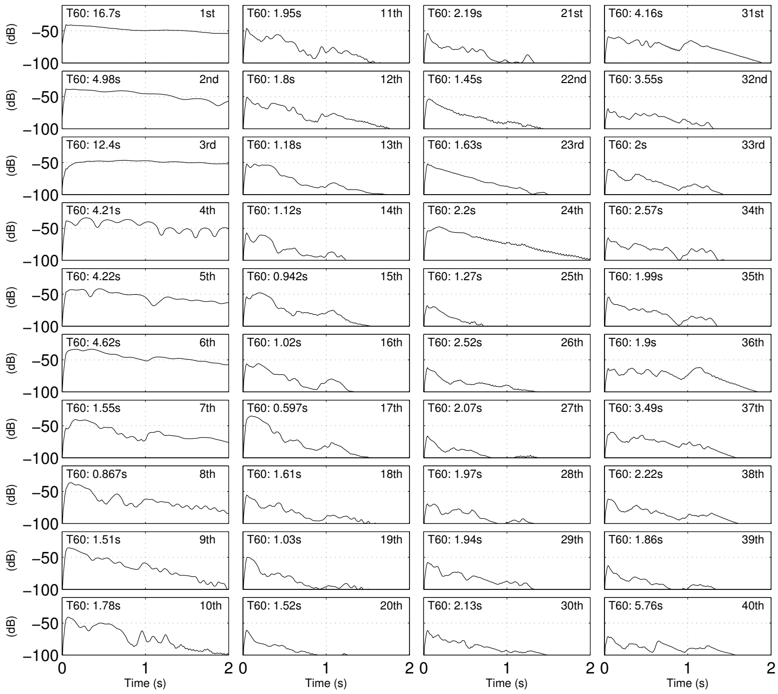

Figure 6.9: Harmonic partials of D2 plucked at one-third of the string

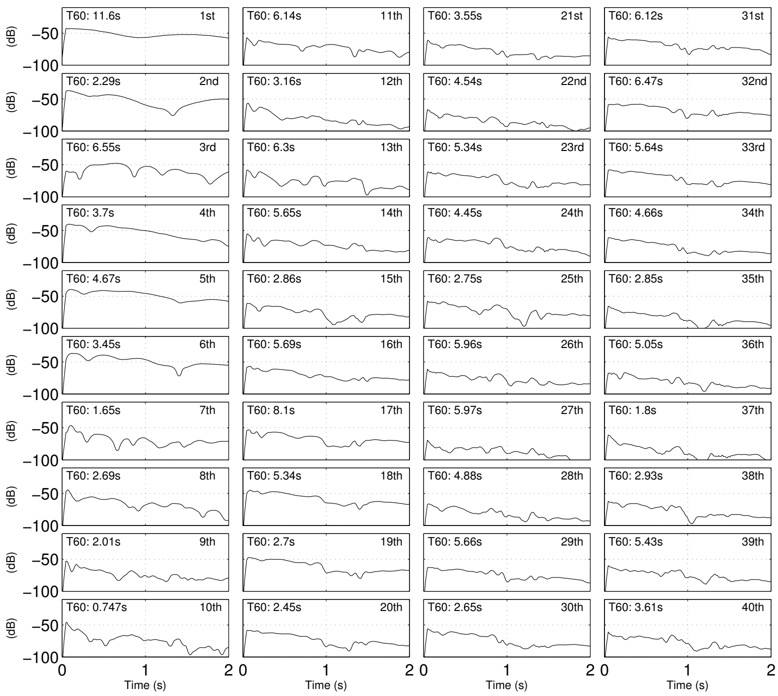

Figure 6.10: Harmonic partials of D3 plucked at one-third of the string

6.3 Spectral Centroid

The spectral centroid is a good measure for the brightness of the sound [57]. It indicates the center of mass of the spectrum. The spectral centroid is calculated as a weighted mean of the frequencies present in the signal. The spectral centroid is plotted as a function of time. The used window size was 2048 samples with step size of 1024 samples.

Figure 6.11 shows spectral centroids of undistorted and distorted string when the string is plucked at the center. The spectral centroids of distorted strings are higher than the one of the undistorted string. The spectral centroid of the undistorted string descends from 7000 Hz to 2000 Hz in half a second. This means that the transient has high frequency content. After one second the spectral centroid has settled at 1000 Hz.

The spectral centroids of the distorted string are higher for longer period than the spectral centroid of the undistorted string. The starting spectral centroid of each distorter is slightly higher, from 500 Hz to 1000 Hz above the spectral centroid of the undistorted string. When the string is plucked at the center D1 and D2 have a similar shape. The spectral centroids of both of the distorters fluctuate which indicates that the amount of distortion changes over time. After 1.5 seconds the higher partials of D1 and D2 die completely. However, D3 maintains its spectral centroid at 4000 Hz for over two seconds. There are moments at the time of one second and 1.5 second in which the spectral centroid is lower meaning that the amount of distortion has decreased. This can also be seen from the spectrogram.

Figure 6.12 presents the spectral centroids of the string as it is plucked at one-third of its length. D1 has a higher spectral centroid than in the case where the string was plucked at the center and the shape resembles more of the spectral centroid of D3. Even after one second the centroid remains above 4000 Hz. This is due to the plucking point. When the string is plucked at one-third of its length there remains more high frequency components as only every third partial is attenuated. The spectral centroid of D2 plucked at one-third of the string is very similar to the plucking at the center. D3 is also similar in both cases. A similar valley as in harmonic partials can be seen at the time from one to 1.2 seconds. This is where the pitch of the partials ascended and descended. During the transition phases the spectral centroid increases and in the middle the spectral centroid is low. The difference of the spectral centroids between the transition phase and the time between the transitions is almost 2000 Hz.

According to the results the distorters increase the spectral centroid of the kantele. The transient is brighter and the distorters maintain the high spectral centroid for a long time period. Because the distorters have increased the brightness of the instrument, the human ear may perceive it louder than without the distorters.

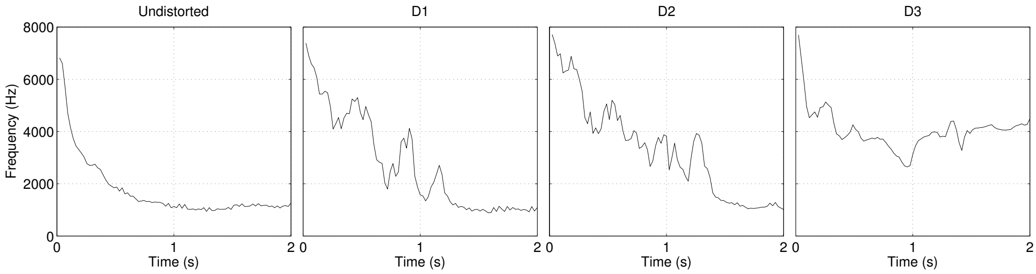

Figure 6.11: Spectral centroids of string plucked at the center of the string

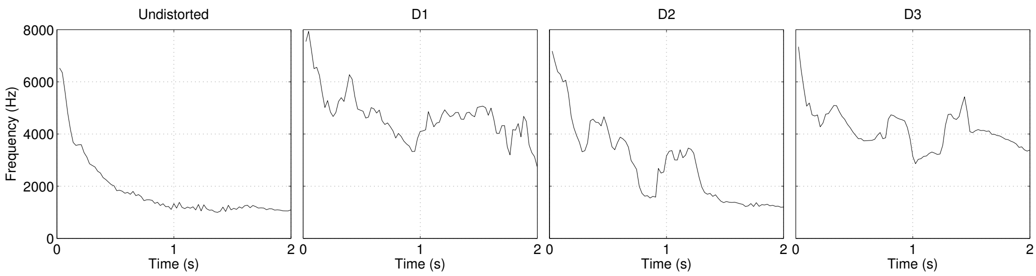

Figure 6.12: Spectral centroids of string plucked at one-third of the string

6.4 A-Weighted Sound Pressure Levels

A-weighting is used in sound level measurements to account for the relative loudness perceived by the human ear. A-weighting filter accents high middle frequencies while attenuating lower frequencies and very high frequencies [18]. It is based on the 30 phon equal-loudness contour. Since the purpose of the distorters is to transfer energy to frequencies to which the human ear is more sensitive, A-weighted signal power levels are a good measure for loudness.

Figure 6.13 shows the original and the A-weighted sound pressure levels when the string is plucked at the center. The original sound pressure level is marked with a solid line while the dashed line presents the A-weighted sound pressure level. The A-weighted sound pressure level of the undistorted signal is lower than the original sound pressure level. After the beginning the difference is up to 2.5 decibels. However, the levels of A-weighted distorted signals are equal to original signals. Since the A-weighting boosts the high middle frequencies and attenuates the low frequencies, this shows that the distorters transfer the energy from the low frequencies to the higher frequencies. Thus, the distorted strings may sound louder than the undistorted string.

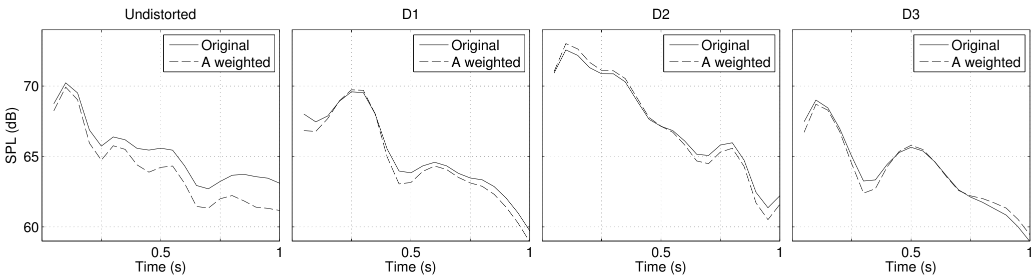

Figure 6.13: Original and A-weighted SPLs of string plucked at one-third of the string

6.5 Discussion

All of the methods state that the distorters increase the brightness of the sound by transferring the energy from the low frequencies to the high frequencies. As the frequency content is spread over a wide band, the harmonic partials cover more critical bands thus making the sound louder. Based on these results the perceived loudness of the instrument increases when a distorter is used.

In addition to loudness, the distorters change the timbre of the string. Since the amount of high frequencies increase, the sharpness of all of the distorters D1, D2 and D3 increases as well. The distorters create a so called buzzing timbre that makes the sound more irritating and increases its audibility. D3 creates a sitar-like timbre that reminds of the Indian stringed instruments. As the timbre changes, the character of the instrument changes. When a distorter is used, the timbre changes so much that the distorted kantele can almost be perceived as a new instrument rather than a traditional kantele.

The amount and the character of the distortion depend on the amplitude of the string. In the beginning, when the vibration amplitude is the highest, the distortion is at it maximum level. When the amplitude decreases, the amount of distortion decreases as well until in the end the string does not touch the distorter anymore and the high harmonic partials vanish. The effect of different plucking forces can easily be tested by plucking the string with a finger. Stronger force creates more distortion. However, if too strong plucking force is applied, the distorters shift from their set position. As a result the string is muted if the contact is too strong or there is no distortion at all, since the string is not in touch with the distorter. This is also one reason why only the 0.1 mm copper wire was used, as it breaks with a smaller force than the wider wires.

The studied distorters are only prototypes and placing of the distorters is problematic. Even a slight change in the position of the distorter changes the timbre of the string. There are endless amount of different timbres between the two extremes, which are an entirely muted string and a free vibrating string. As there is no best type of distortion, the best setup depends on the personal taste. The designer Jyrki Pölkki stated that the distorters are meant to be used in a way that they sound the best.

Scientifically the problem is that it is almost impossible to setup distorters similarly as before. Every pluck of the string is unique, which is why only one sample of each distorter was selected for the results. Each plucking of the string also moves the distorter from its place. Sometimes this made distortion sound better and sometimes the distorter did not even touch the string anymore. The stronger the plucking force was, the more it could displace the distorter from its optimal place thus ruining it. This complicated the measurements, since every time was different.

There exist some problems that have to be solved before the product can be launched commercially. For a professional player it is crucial that the instrument they are playing is reliable. If the distorter creates a different distortion on each pluck, the distorter is unusable for actual playing. The player should be able to lock the distorter to its position. Currently, finding of a good distortion depends mainly on luck. It is difficult to set the distorter in the same way as before.

While it is important that the distortion maintains its setup, there should be also a way to alter it for different timbres. The guitar effect pedal manufacturers have solved this problem with knobs that can be used to change predetermined values. This way the user can adjust the timbre in the limits of the parameters. The mechanical distorters would benefit of similar adjustment tools. Fine-tuning of the studied distorters is very hard. However, fine-tuning of most of the Indian stringed instruments is so hard that it is mostly done by professional instument builders.

Chapter 7

Conclusions

The traditional Western stringed instruments are designed to produce sounds that have most of their frequency components in the low end of the spectrum. They perform well in a quiet environment, but get masked under background noise in noisy environments. This can be overcome by increasing the volume of the instrument. However, since the human ear is the most sensitive for high middle frequencies and broadband sounds, the timbre can be changed to increase the perceived loudness.

A mechanical distorter transfers energy from the low frequencies to the higher frequencies thus making the sound more broadband. This makes the instrument louder and helps it cut through the surrounding noise. A distorter can be a piece of metal or rawhide that is attached close to the string. As the string vibrates the distorter limits the vibrating amplitude thus creating nonlinear distortion. According to the studies a mechanical distorter transfers energy from low partials to high partials and extends the duration of the high components. However, the distorters change the character of the instrument making it sound like a completely new instrument.

This thesis studies three different mechanical distorters which are made of a curved metal wire or rawhide. The distorters were placed close to the termination of the string near the tuning pin of the kantele. As the string vibrates it collides with the distorters. The audio samples of the undistorted string and the three distorted strings were compared.

The measured audio samples were studied by using four different methods: the spectrograms, the harmonic partials, the temporal evolution of the spectral centroid, and the A-weighted sound pressure levels. The spectrograms show that each of the distorters enhances the energy and the duration of the high frequencies. The isolated harmonic partials give similar results. The strings with a distorter have a higher spectral centroid than the one without. Thus the distorters make the tone brighter than the original. A-weighted sound pressure level is used to approximate the perceived loudness of the sound. A-weighted SPL of the undistorted string is lower than the original SPL. However, the A-weighted and the original sound pressure levels of the distorted strings are equally loud. Based on comparison of the standard SPL and the A-weighted SPL of the tones the distorted strings are perceived louder.

Bibliography

[1] T. D. Rossing, *The Science of String Instruments.* Springer, 2010.

[2] H. Penttinen, C. Erkut, J. Pölkki, V. Välimäki, and M. Karjalainen, "Design and analysis of a modified kantele with increased loudness," *Acta Acustica united with Acustica*, vol. 91, pp. 261–268, 2005.

[3] H. Tahvanainen, "Modelling Body Vibration and Sound Radiation of a Modified Kantele," Master's thesis, Aalto University School of Electrical Engineering, 2012.

[4] H. Tahvanainen, J. Pölkki, H. Penttinen, and V. Välimäki, "Finite Element Model of the Kantele with Improved Sound Radiation," in *Proceedings of the Stockholm Music Acoustic Conference, Stockholm, Sweden*, pp. 193–198, 2013.

[5] M. Schatzman, "A Hyperbolic Problem of Second Order with Unilateral Constraints: The Vibrating String with a Concave Obstacle," *Journal of Mathematical Analyses and Applications*, vol. 73, pp. 138–191, 1980.

[6] J. K. Robert Burridge and C. Morshedi, "The Sitar String, a Vibrating String with a One-Sided Inelastic Constraint," *SIAM Journal on Applied Mathematics*, vol. 42, pp. 1231–1251, 1982.

[7] K. A. Legge and N. H. Fletcher, "Nonlinear generation of missing modes on a vibrating string," *Journal of the Acoustical Society of America*, vol. 76, no. 1, pp. 5–12, 1984.

[8] M. Frontini and L. Gotusso, "Numerical study of the motion of a string vibrating against an obstacle by physical discretization," *Applied Mathematical Modelling*, vol. 14, pp. 489–494, 1990.

[9] E. Rank and G. Kubin, "A waveguide model for slapbass synthesis," in *Proceedings of 1997 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 1, pp. 443 – 446, 1997.

[10] A. Krishnaswamy and J. O. Smith, "Methods for Simulating String Collisions with Rigid Spatial Obstacles," in *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pp. 233–236, 2003.

[11] G. Evangelista and F. Eckerholm, "Player-Instrument Interaction Models for Digital Waveguide Synthesis of Guitar: Touch and Collisions," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, pp. 822–832, 2010.

[12] T. Taguti, "Dynamics of simple string subject to unilateral constraint: A model analysis of sawari mechanism," *Acoustical Science and Technology*, pp. 203–214, 2008.

[13] H.-M. Lehtonen, A. Askenfelt, and V. Välimäki, "Analysis of the part-pedaling effect in the piano," *Journal of the Acoustical Society of America*, vol. 126, pp. EL49–EL54, 2009.

[14] A. Stulov, "Vibration of strings with nonlinear supports," *Applied Acoustics*, vol. 76, pp. 223–229, 2014.

[15] S. Bilbao and A. Torin, "Numerical Simulation of String/Barrier Collisions: The Fretboard," in *Proceedings of the 17th International Conference on Digital Audio Effects (DAFX-14), Erlangen, Germany*, 2014.

[16] M. Karjalainen, J. Backman, and J. Pölkki, "Analysis, Modeling, and Real-time Sound Synthesis of the Kantele, a Traditional Finnish String Instrument," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, (Minneapolis, Minnesota, USA), pp. 229–232, 1993.

[17] C. Valette, "The mechanics of vibrating strings," in *Mechanics of Musical Instruments* (A. Hirschberg, J. Kergomard, and G. Weinreich, eds.), pp. 115–184, Springer, 1995.

[18] T. D. Rossing, R. F. Moore, and P. A. Wheeler, *Science of Sound*. Addison Wesley, 3rd ed., 2001.

[19] N. H. Fletcher and T. D. Rossing, *The Physics of Musical Instruments*. Springer Verlag, 2nd ed., 1998.

[20] H. Penttinen, *Loudness and Timbre Issues in Plucked Stringed Instruments – Analysis, Synthesis, and Design*. PhD thesis, Helsinki University of Technology, Espoo, Finland, 2006.

[21] Harold A. Conklin Jr, "Piano design factors - their influence on tone and acoustical performance." `https://www.speech.kth.se/music/5_lectures/conklin/longitudinal.html`, November.

[22] B. Bank and L. Sujbert, "Modeling the Longitudial Vibration of Piano Strings," in *Proceedings of the Stockholm Music Acoustics Conference (SMAC 03), Stockholm, Sweden*, 2003.

[23] B. Bank and L. Sujbert, "Generation of longitudinal vibrations in piano strings: From physics to sound synthesis," *Journal of the Acoustical Society of America*, vol. 117, pp. 2268–2278, 2005.

[24] M. Demoucron, S. Weisser, and M. Leman, "Sculpting the sound. timbre-shapers in classical Hindustani chordphones," in *Proceedings of the 2nd CompMusic Workshop, Istanbul, Turkey*, 2012.

[25] S. Siddiq, "Physical modelling of the sitar string," in *Proceedings of the Second Vienna Talk, Vienna, Austria*, pp. 137–140, 2010.

[26] C. Erkut, *Aspects in Analysis and Model-based Sound Synthesis of Plucked String Instruments*. PhD thesis, Helsinki University of Technology, Espoo, Finland, 2002.

[27] W. A. Sethares, *Tuning, Timbre, Spectrum, Scale*. Springer Verlag, 1st ed., 1997.

[28] M. Karjalainen, *Kommunikaatioakustiikka*. Helsinki University of Technology, 1st ed., 1999.

[29] B. C. J. Moore, *An Introduction to the Psychology of Hearing*. Academic Press, 4th ed., 1997.

[30] Castle Group. `http://www.castlegroup.co.uk/frequency-responses/`, November 2014.

[31] J. M. Grey, *An Exploration of Musical Timbre*. PhD thesis, Stanford University, Stanford, United States, 1975.

[32] G. Davis and R. Jones, *The Sound Reinforcement Handbook*. Yamaha, 2nd ed., 1988.

[33] J. J. Fricke, "Jimi Hendrix' Use of Distortion to Extent the Performance Vocabulary of the Electric Guitar," *Journal of the Acoustical Society of America*, vol. 103, pp. 443–444, 1998.

[34] M. Atsushi and W. L. Martens, “Timbre of Nonlinear Distortion Effects: Perceptual Attributes Beyond Sharpness?,” in *Proceedings of the Conference on Interdisciplinary Musicology (CIM05), Montreal, Canada*, pp. 101–138, 2005.

[35] D. J. Dailey, *Electronics for Guitarists*. Springer Science & Business Media, 2011.

[36] A. Adler, V. Emiya, M. G. Jafari, M. Elad, R. Gribonval, and M. D. Plumbley, “A Constrained Matching Pursuit Approach to Audio Declipping,” in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing, Prague, Czech Republic*, pp. 329–332, 2011.

[37] F. Hoffmann, *Encyclopedia of Recorded Sound*. Routledge, 2nd ed., 2004.

[38] J. Pakarinen and D. T. Yeh, “A Review of Digital Techniques for Modeling Vacuum-Tube Guitar Amplifiers,” *Computer Music Journal*, vol. 33, pp. 85–100, 2009.

[39] R. Denyer, *The Guitar Handbook*. Knopf, 7th ed., 1992.

[40] C. V. Raman, “On Some Indian Stringed Instruments,” in *Proceedings of the Indian Association for the Cultivation of Science*, vol. 7, pp. 29–33, 1922.

[41] T. H. Park and Z. Li, “All this buzz about the sitar: Physically modelingly speaking,” in *Proceedings of the 15th International Congress on Sound and Vibration, Daejeon, Korea*, pp. 907–913, 2008.

[42] T. Taguti, “Sound production of sawari mechanism in biwa,” *Acoustical Science and Technology for Quality of Life*, pp. 2119–2122, 2004.

[43] S. Yoshikawa, “Psaltiers and zithers,” in *The Science of String Instruments* (T. D. Rossing, ed.), pp. 173–196, Springer, 2010.

[44] D. Ronan, “The physical modelling of a sitar,” Master’s thesis, University of Limerick, Limerick, Ireland, 2010.

[45] Bachido: International Shamisen Community. `http://bachido.com/overview/shamisen-styles`, November 2014.

[46] T. Taguti and Y. K. Tohnai, “Acoustical analysis on the sawari tone of chikuzen biwa,” *Acoustical Science and Technology*, pp. 199–207, 2001.

[47] S. Weisser and O. Lartillot, "Investigating non-western musical timbre: A need for joint approaches," in *Proceedings of the Third International Workshop on Folk Music Analysis, Amsterdam, Netherlands*, pp. 33–39, 2013.

[48] D. Kartofelev, A. Stulov, H.-M. Lehtonen, and V. Välimäki, "Modeling a vibrating string terminated against a bridge with arbitrary geometry," in *Proceedings of the Stockholm Music Acoustics Conference 2013: 4th Stockholm Music Acoustics Conference, Stockholm, Sweden*, pp. 626–632, 2013.

[49] M. V. Walstijn and V. Chatziioannou, "Numerical Simulation of Tanpura String Vibrations," in *ISMA 2014, Le Mans, France*, pp. 609–614, 2014.

[50] A. Peekna and T. D. Rossing, "Psaltiers and zithers," in *The Science of String Instruments* (T. D. Rossing, ed.), pp. 99–122, Springer, 2010.

[51] C. Erkut, M. Karjalainen, P. Huang, and V. Välimäki, "Acoustical Analysis and Model-based Sound Synthesis of the Kantele," *Journal of the Acoustical Society of America*, vol. 112, no. 4, pp. 1681–1691, 2002.

[52] "Rode NT1-A Datasheet." `www.rodemic.com/download/nt1-a_datasheet.pdf`, November 2014.

[53] J. Woodhouse, "Plucked Guitar Transients: Comparison of Measurements and Synthesis," *Acta Acustica United with Acustica*, vol. 90, pp. 945–965, 2004.

[54] J. Rinne, J. Pölkki, H. Tahvanainen, and V. Välimäki, "Mechanical Distortion in Plucked String Instruments," in *Baltic-Nordic Acoustic Meeting (BNAM 2014), Tallinn, Estonia*, pp. 447–453, 2014.

[55] J. M. Grey and J. W. Gordon, "Perceptual Effects of Spectral Modifications on Musical Timbres," *Journal of the Acoustical Society of America*, vol. 63, pp. 1270–1277, 1978.

[56] H.-M. Lehtonen, V. Välimäki, and T. I. Laakso, "Canceling and Selecting Partials From Musical Tones Using Fractional-Delay Filters," *Computer Music Journal*, vol. 32, pp. 43–56, 2008.

[57] B. C. J. Moore, *Hearing.* Academic Press, 1995.